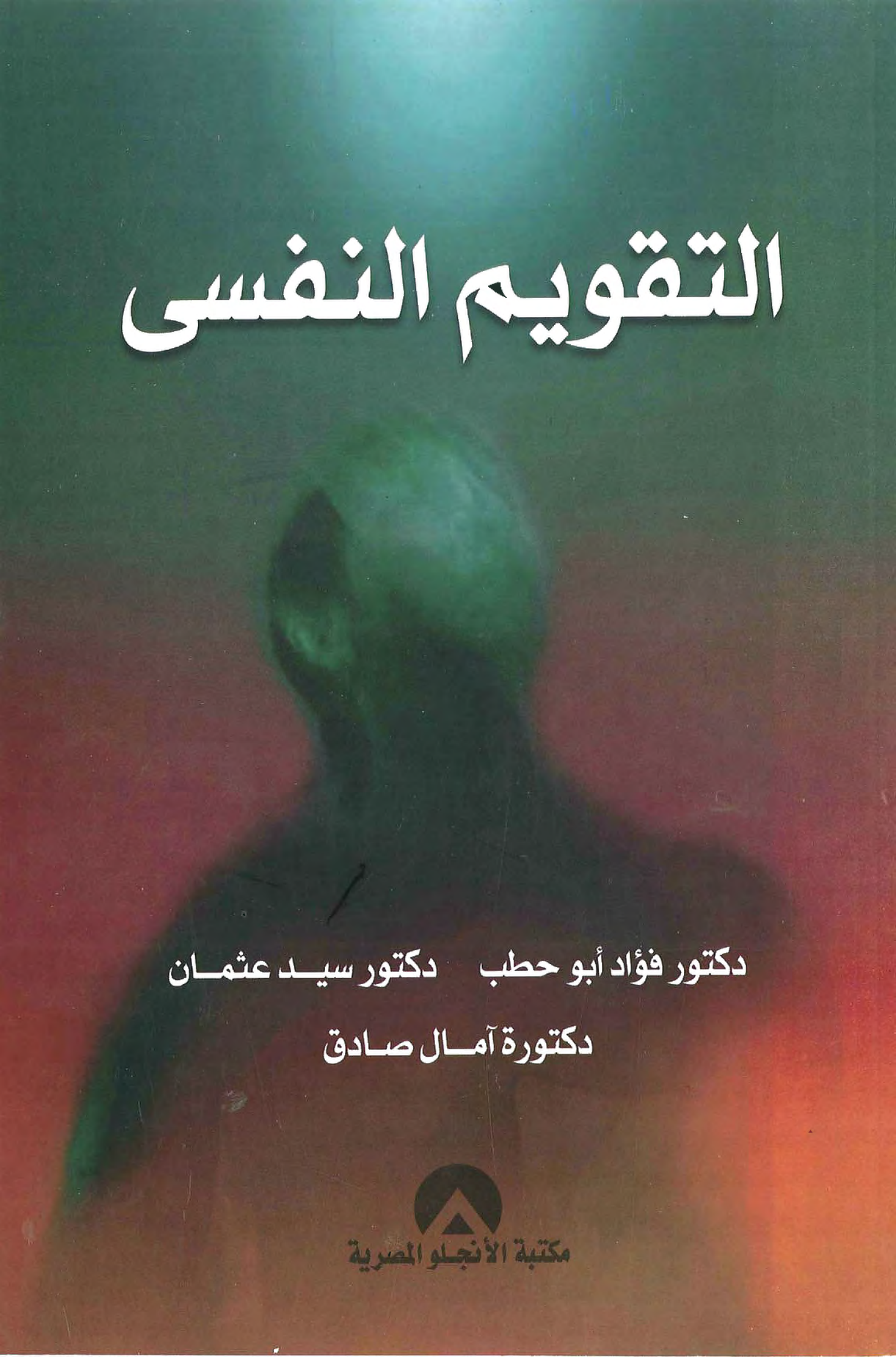
التقويم النفسى



دكتور فؤاد أبو حطب دكتور سيد عثمان

دكتورة آمال صادق

مكتبة الأنجلو المصرية

التقويم النفسى

تأليف

دكتور فؤاد أبو حطب
أستاذ علم النفس
قسم علم النفس التعليمى
كلية التربية جامعة عين شمس

دكتور سيد أحمد عثمان
أستاذ علم النفس
قسم علم النفس التعليمى
كلية التربية جامعة عين شمس

دكتورة آمال صادق
أستاذ علم النفس
ونائب رئيس جامعة حلوان
(الأسبق)

الطبعة الرابعة
(مزيده ومجددة)

مكتبة الأنجلو المصرية

صدر هذا الكتاب لأول مرة بعنوان (مشكلات التقويم النفسى) عام ١٩٧٠ ثم صدر بعنوان (التقويم النفسى) كما يلى:

الطبعة الأولى ١٩٧٣

الطبعة الثانية ١٩٧٦

إعادة طبع ١٩٧٩

إعادة طبع ١٩٨٣

إعادة طبع ١٩٨٤

إعادة طبع ١٩٨٥

الطبعة الثالثة ١٩٨٦

إعادة طبع ١٩٩٧

الطبعة الرابعة ٢٠٠٨

جميع حقوق الطبع محفوظة للمؤلفين

ولا يجوز نقل أى فقرة دون الإحالة للمصدر

رقم الإيداع ٢٣١١

الرقم الدولى ٣ – ٠٥٢٩ – ٠٥ – ٩٧٧

تقديم الطبعة الرابعة

عندما صدر كتاب (مشكلات فى التقويم النفسى) فى مطلع عام ١٩٧٠ . والذى ألفه كل من أ.د. فؤاد أبو حطب وأ.د. سيد أحمد عثمان لم يكن التوقع له أن يتطور وينمو على النحو الذى صار إليه، لقد استقبل هذا الكتاب يومئذ من دارسى علم النفس فى مصر والوطن العربى استقبالا ابهج مؤلفاه وقررا أن يوليا ميدان التقويم النفسى مزيدا من الاهتمام، ولم تنقض سنوات ثلاث حتى نفد هذا الكتاب الصغير وصدرت فى عام ١٩٧٣ الطبعة الأولى من كتاب (التقويم النفسى) فى صورته الأكثر شمولا واتساعا وتقدما من الكتاب السابق. ولم تمر سنوات ثلاث أخرى إلا وصدرت فى عام ١٩٧٦ الطبعة الثانية من هذا الكتاب تضيف إلى الطبعة الأولى وتصحح بعض ما جاء فيها.

وشاءت ظروفنا أن تمضى السنوات العشر التالية دون أن تصدر طبعة جديدة من هذا الكتاب، وكل ما نشر فى أعوام ١٩٧٩، ١٩٨٣، ١٩٨٤، ١٩٨٥ لم يكن سوى إعادة إصدار للطبعة الثانية، بعضه ظهر بعلمنا وعن ناشرنا، وبعضه الآخر كان تصويرا وتزويرا للكتاب قام به ناشرون آخرون، وظاهرة تصوير الكتب وتزويرها أصبحت جزءا خطيرا من أزمة النشر التى يتعرض لها الكتاب المصرى فى السنوات الأخيرة، زاد من سعارها إقبال القارئ على الكتاب الجيد، وعلى الرغم من أن هذا يتضمن فى ذاته حكما بالجودة على كتابنا إلا أنه فى الوقت نفسه يحمل نذر خطر أخلاقى يهدد بالشر المستطير.

كان الانقطاع عن تجديد الكتاب طوال السنوات العشر الماضية يسبب لنا آلاما مضنية. فقضايا التقويم النفسى «متعددة متجددة» كما وصفنا منذ صدر كتاب (مشكلات فى التقويم النفسى). وقد حددنا منذ ذلك الحين مظاهر هذا التعدد والتجدد وحصرناها فى ثلاث جوانب: أولاها طبيعة الظواهر النفسية موضع التقويم، وثانيتها طبيعة أدوات التقويم النفسى، وثالثتها طبيعة المشكلات التى تنشأ عن استخدام هذه الأدوات فى مجالات التطبيق.

والحق أن هذه الجوانب الثلاثة شهدت - منذ منتصف السبعينات - ولاتزال تشهد تطورات وتغيرات هامة يكاد يصل بعضها إلى حد «الثورة العلمية» الكاملة ولعل أكثر مظاهر هذا التطور عمقا ما طرأ على النظرية السيكومترية العامة من اتساع لآفاقها وامتداد لحدودها جعلها تتجاوز «الاختبارات المرجعة إلى المعيار» إلى «الاختبارات

المرجعة إلى المحك»، وعلى نحو تطلب إعادة النظر فى كثير من المفاهيم الكلاسيكية لنظرية القياس النفسى.

أما أكثر مظاهر التطور حدة فى ميدان التقويم النفسى فهو ما يشهده مجال التطبيق من قضايا خلافية ظهرت على وجه الخصوص فى المجتمعات التى تعد الاختبارات النفسية جزءا من بنيتها المعرفية العامة، وعلى وجه أخص بريطانيا والولايات المتحدة الأمريكية، وكان لهذا الجدل الحاد تضميناته الاجتماعية والثقافية والأيديولوجية.

يضاف إلى ذلك ما يمكن أن نسميه «أزمة الاختبارات النفسية فى مصر والوطن العربى». وهى أزمة معقدة متعددة الجوانب تشمل فيما تشمل ندرة الإبداع فى هذا المجال السيكولوجى الهام، والاعتماد فى كثير مما ظهر من الاختبارات النفسية على النقل المباشر أو الترجمة الحرفية أو المحرفة عن اللغات الأوربية - وخاصة اللغة الإنجليزية، ناهيك عن النقص المتزايد فى أعداد المتخصصين المتمرسين فى هذا الميدان ، مع زيادة مطردة فى اعداد الهواة.

هذا بالإضافة إلى وجود العديد من أدوات التقويم والقياس تزخر بها دور النشر المعروفة أو غير المعروفة مما يهدد بفقدان هذه الأدوات فعاليتها وموثوقيتها لاتاحتها إلى كل من يستطيع شراءها.

ونقترح للتغلب على هذه المشكلة ان تتولى هيئة علمية مسئولة عن نشر الإختبارات السيكولوجية أو التصريح بنشرها حيث ان الكثير منها يفتقر إلى مقومات الإختبار الجيد وشروطه، وربما يكون هذا الاقتراح تحديدا أن تتولى الجمعية المصرية للدراسات النفسية هذه المهمة أسوة بما هو معمول به فى الخارج حيث توجد هيئات علمية مسئولة عن نشر الإختبارات مثل NFER فى المملكة المتحدة.

ولعل أكثر ما يدعو إلى الضيق حقا أن التقويم النفسى الذى يمتد عمره الزمنى فى مصر إلي أكثر من خمسين عاما لم يتجاوز بعد الأروقة الأكاديمية الضيقة، وهو وضع نجد ما يشبهه فى جميع الأقطار العربية، ولعل هذا يدفعنا إلى طرح تساؤل ربما يكون موضعا لبحث علمى جاد: هل أسباب ذلك فينا نحن السيكولوجيين الذين عجزنا عن أن نصنع قنوات اتصال جيدة مع ثقافتنا؟ أم أنها فى مجتمعاتنا التى لاتزال تتجنب العلم سبيلا إلى حل مشكلاتها؟ أم هى فينا وفى مجتمعاتنا معا.

كان هذا كله وغيره عند إصدار الطبعة الثالثة من هذا الكتاب واستمر نفس الإهتمام عند تحديث هذا الكتاب أيضا فى طبعته الرابعة. فالكتاب يقع فى ستة أبواب حيث جاء الباب الأول فى فصلين تناول الأول منها متضمنا الأسس العامة للتقويم النفسى ومفاهيمه ومهامه وتطوره وأدواره والقضايا الخلافية المعاصرة، أما الفصل الثانى فيقدم للقارىء المبتدىء مدخلا مبسطا للإحصاء النفسى فى ابسط صورة، وربما يمكن الاستغناء عنه للقارئ المتخصص والذى يمكنه استخدام الحاسبات والبرامج الإحصائية والتى يسرت على الباحثين عناء التحليلات الاحصائية اللازمة فى اغراض التقويم والقياس.

أما الباب الثانى فيعرض لمبادىء القياس النفسى، وجاء فى أربعة فصول تناولت على التوالى موضوعات: فن إعطاء الاختبارات، والثبات، والصدق، وتصحيح الاختبارات وتفسير الدرجات، وقد تناول هذا الباب معظم الجديد فى النظرية السيكومترية العامة والذى شهدته طوال العقد المنقضى.

وقد خصص الباب الثالث «لاختبارات الذكاء العام» ويقسم إلى فصلين أحدهما حول اختبارات الذكاء الفردية وعلى وجه أخص اختبارات ستانفورد - بينيه ووكسلر، وثانيهما حول اختبارات الذكاء الجماعية مع تناول مفصل لهذه الاختبارات بنوعيها والبحوث المصرية والعربية التى أجريت حولها.

أما الباب الرابع فحول بطاريات الاستعدادات ويقسم أيضاً إلى فصلين أحدهما حول بطاريات الاستعدادات المتعددة، وثانيهما حول اختبارات القدرات المتخصصة وقد عرضنا هذه الاختبارات بالتفصيل الكافى مع تناول الجهود المصرية والعربية التى بذلت فى بناء بعض هذه الاختبارات أو فى تقنينها.

وأفردنا فى هذه الطبعة الباب الخامس لاختبارات التحصيل، وجاء هذا الباب فى فصلين تناولت أولهما بناء الاختبار التحصيلى، أما ثانيهما فحول الاختبارات المرجعة إلى المحك، وهو فصل جديد يتناول هذا الاتجاه المهم فى القياس التربوى.

أما الباب السادس فيتناول مقاييس الشخصية، ويتألف من أربعة فصول تناولت على التوالى: وسائل التقرير الذاتى، استخدام الأحكام والملاحظات المنظمة، استخدام الأداء والأساليب الإسقاطية، وقياس الميول.

وختمنا كل باب من أبواب الكتاب الستة بمراجعه، ثم ختمنا الكتاب كله بقائمة

مركزة تساعد المشتغل بالتقويم النفسى على تقويم أدواته التى يستعملها.

أما هذه الطبعة التى بين أيدى القارىء فصدورها يأتى بعد وفاة الراحل العظيم فؤاد أبو حطب مؤلف الكتاب الأول ورائد القياس النفسى فى مصر. والذى أضاف الكثير فى مجال التقويم النفسى والتى تزخر هذه الطبعة بما أضافه من كتابات تتضمن ما أنجزه منذ صدور الطبعة السالفة بعد صدور الطبعة وحتي وفاته ولم يتم نشرها.

وقد آثرت أن تعرضها فى هذه الطبعة أ.د. آمال أحمد مختار صادق رفيق عمره، كما تم اضافة ما تم تحت اشرافه في مجال القياس والتقويم التربوى كمدير للمركز القومى للإمتحانات والتقويم التربوى وما تناولته المؤتمرات والندوات العلمية والاوراق البحثية والتى ترتبط ارتباطا مباشرا بموضوعات هذا الكتاب.

هذا بالإضافة إلى عرض بعض الجهود التى أضافها أصحابها فى مجال القياس النفسى والتقويم التربوى سواء بإعادة تقنين بعض الإختبارات والمقاييس، أو بناء الجديد منها للنهوض بهذا المجال، ومسايرة المستحدثات التى ارتبطت بتطوير نظام التعليم فى مصر وما يتطلبه هذا العصر من تطوير فى نظم الإنتقاء والتصنيف والتوجيه والإرشاد التعليمى وهذا لا يتم بالطبع إلا من خلال الاهتمام بمجال التقويم والقياس النفسى والتربوى وهو بالطبع ما نصبو له.

وقد آثرنا الإحتفاظ بالشكل العام للكتاب وتصنيف فصوله، أما التحديث فارتبط بما تم اضافته من معلومات فى كل فصول الكتاب وجاءت خاتمته بما سطره أ.د. فؤاد أبو حطب عن تقويم التقويم وتقويم أدوات التقويم.

وفقنا الله إلى ما فيه الخير والله ولى التوفيق.

القاهرة

يوليو ٢٠٠٨

المؤلفون

التقويم النفسى ـــــــ

تحليل بيانات مقاييس الرتبة ٩٧
الوسيط ٩٧
نصف المدى الربيعى ٩٩
معامل ارتباط الرتب ١٠٠
تحليل بيانات مقاييس المسافة والنسبة ١٠٢
المتوسط الحسابى ١٠٢
الانحراف المعيارى ١٠٣
معامل الارتباط التتابعى ١٠٤
مراجع الباب الأول ١٠٥
الباب الثانى: شروط الاختبارات النفسية ١١١
الفصل الثالث: فن إعطاء الاختبارات ١١٣
ماذا يقصد بتقنين الاختبار؟ ١١٣
واجبات الفاحص ومسئولياته ١١٤
طريقة إعطاء الاختبار ١١٧
استثارة الدوافع ١٢٢
التهيؤ لأخذ الاختبار ١٢٥
عملية الاختبار باعتبارها علاقة اجتماعية ١٢٧
دور اللغة فى تكوين العلاقة الاجتماعية ١٣٠
كيف يشوه الفاحصون نتائج الاختبارات ١٣٢
الفصل الرابع: ثبات الاختبارت ١٣٥
معنى الثبات ١٣٥
مصادر التباين ١٣٧
طرق حساب الثبات ١٤١
إعادة الاختبار ١٤١
الصور المتكافئة أو البديلة ١٤٤
التجزئة التصفية ١٤٦
طريقة كيودر– ريتشاردسون ١٥٠
معامل ألفا لكرونباك ١٥٢
ثبات الفاحصين والمصححين ١٥٣

العوامل المؤثرة فى ثبات الاختبار ١٥٥
الخطأ المعيارى للقياس ١٦٠
الفصل الخامس: صدق الاختبارات ١٦٥
معنى الصدق ١٦٥
طرق تحديد الصدق ١٦٥
صدق المحتوى ١٦٥
طرق تحديد صدق المحتوى ١٦٨
جدول المواصفات ١٦٨
بعض الطرق الإمبريقية لتحديد صدق المحتوى ١٧٤
الصدق المرتبط بالمحكات ١٨١
مشكلات الصدق المرتبط بالمحكات ١٨٩
صدق التكوين الفرضى ١٩٠
العوامل المؤثرة فى الصدق ١٩٩
الفصل السادس: تصحيح الاختبارات وتفسير الدرجات ٢٠٣
تصحيح الاختبار ٢٠٣
تصحيح الاختبارات ذات الاستجابة الحرة ٢٠٤
تصحيح اختبارات الاختيار من متعدد ٢٠٦
تصحيح أثر التخمين ٢٠٧
تفسير الدرجات الخام ٢١٠
المعايير ٢١١
المعايير الارتقائية أو النمائية ٢١٢
نسبة الذكاء التقليدية ٢١٤
المئينيات ٢١٦
الدرجات المعيارية ٢١٨
المحكات ٢٢٣
جداول التوقع ٢٢٣
المستويات ٢٢٦
مستوى نطاق المحتوى ٢٢٦
المستويات الارتقائية الرتبية ٢٢٧

مراجع الباب الثانى ٢٢٩
الباب الثالث: اختبارات الذكاء العام ٢٣٣
تقديم ٢٣٥
الفصل السابع: اختبارات الذكاء الفردية ٢٣٩
مقياس ستانفورد ـ بينيه ٢٣٩
نشأة المقياس وتطوره ٢٤٠
تحليل أسئلة المقياس ٢٥٤
وصف المقياس ٢٥٦
تطبيق المقياس ٢٥٨
تصحيح المقياس ٢٥٩
تفسير المعايير ٢٦٦
ثبات المقياس ٢٦٩
صدق المقياس ٢٧١
الطبعة الخامسة من المقياس ٢٧٨
مقاييس وكسلر للذكاء ٢٨١
مقياس وكسلر لذكاء الراشدين ٢٨٤
مقياس وكسلر لذكاء الأطفال ٢٩٣
مقياس وكسلر لذكاء أطفال ما قبل المدرسة ٢٩٧
الصور المختصرة من مقاييس وكسلر ٢٩٩
المقياس البريطانى للقدرة ٢٩٩
مقاييس القدرات الفارقة ٣٠٣
منظومة التقييم المعرفى ٣٠٦
بطارية كوفمان ٣٠٧
مقياس كوفمان لذكاء الاطفال ٣٠٩
مقياس كوفمان للمراهقين والكبار ٣٠٩
البطارية المختصرة لكوفمان ٣١٠
قياس ذكاء الرضع والاطفال ٣١٠
– الجداول الارتقائية لجزل ٣١١
خفض مستوى ستانفورد – بينيه ٣١١

مقاييس بيلى لنمو الرضع ٣١٣
مقاييس مكارثى لقدرات الاطفال ٣١٤
الإتجاه النمائى فى قياس وتفسير الذكاء ٣١٦
المقاييس المعتمدة على نموذج بياجيه ٣١٦
الاتجاهات المعاصرة فى تقدير سلوك الاطفال فى مرحلة الطفولة المبكرة... ٣١٩
جهود المركز القومى للامتحانات والتقويم التربوي فى تحديد معايير نمو... ٣٢١
الأطفال
الاختبارات العملية للذكاء ٣٢٢
الفصل الثامن: الاختبارات الجماعية للذكاء ٣٢٧
أهمية الاختبارات الجماعية ٣٢٧
خصائص الاختبارات الجماعية ٣٢٧
دور الحاسب الألكترونى ٣٢٩
نظرية السمة الكامنة أو الاستجابة على المفردة ٣٣٠
بطاريات اختبارات الذكاء المتعددة المستويات ٣٣١
بطاريات أطفال الحضانة ورياض الأطفال ٣٣٣
بطاريات تلاميذ مرحلة التعليم الأساسى ٣٣٥
اختبارات ذكاء طلاب المرحلة الثانوية والجامعات ٣٤١
اختبارات ذكاء طلاب الدراسات العليا ٣٤٤
اختبارات ذكاء الراشدين المتفوقين ٣٤٥
اختبارات الذكاء المتحررة من أثر الثقافة ٣٤٥
اختبارات كاتل غير المتحيزة للثقافة ٣٤٧
اختبارات المصفوفات المتتابعة ٣٤٩
اختبار رسم الرجل ٣٥٥
نحو مفهوم شامل للذكاء ٣٦٢
تضمينات النموذج الرباعى المعلوماتى عند فؤاد ابو حطب ٣٦٢
مراجع الباب الثالث ٣٨٥
الباب الرابع: مقاييس الاستعدادات ٣٩٤
الفصل التاسع: بطاريات الاستعدادات المتعدد ٣٩٧
اختبارات القدرات العقلية الأولية ٣٩٨

اختبارات الاستعدادات الفارقة ٤٠٤
بطارية الاستعدادات العامة ٤٠٧
اختبارات فلانجان لتصنيف الاستعدادات ٤٠٩
الاختبارات العملية ٤١٣
بطارية اختبارات استعدادات حرف المعادن ٤١٨
بطارية الاستعدادات الفارقة لمورسبى ٤٢١
بطارية الاختبارات المعرفية مرجعية العوامل ٤٢٢
الفصل العاشر: اختبارات القدرات المتخصصة ٤٢٥
اختبارات القدرات الحركية ٤٢٦
اختبارات القدرات الميكانيكية ٤٢٧
اختبارات القدرة الكتابية ٤٢٩
اختبارات القدرات الموسيقية ٤٣٢
اختبارات القدرة فى الفنون البصرية ٤٤٠
اختبارات التفكير الابتكارى ٤٤٢
مقياس التفكير الأخلاقى للأطفال ٤٥٤
اختبارات التفكير الناقد ٤٥٥
بطارية الاستعدادات الحسية الحركية للمكفوفين ٤٥٧
اختبارات التذكر ٤٥٧
اختبار الأشكال المتضمنة ٤٥٨
مراجع الباب الرابع ٤٥٩
الباب الخامس: اختبارات التحصيل ٤٦٥
الفصل الحادى عشر: بناء الاختبار التحصيلى ٤٦٧
تحديد الأهداف التعليمية ٤٦٧
وصف العمل التربوى ٤٦٨
تحليل العمل التربوى ٤٦٩
وصف المدخلات السلوكية ٤٧٠
العلاقة بين الأهداف التعليمية وأساليب التقويم ٤٧١
تحليل محتوى المادة الدراسية ٤٧٤
إعداد جدول المواصفات ٤٧٥

تحديد نوع الأسئلة ٤٧٦
الاختبارات الشفوية ٤٨٠
تقويم التحصيل فى المدرسة الابتدائية ٤٨٥
تقويم التحصيل فى القراءة ٤٨٥
تقويم التحصيل فى الحساب ٤٨٧
تقويم التحصيل فى العلوم ٤٨٩
الفصل الثانى عشر: الاختبارات المرجعة إلى المحك ٤٩٣
مقدمة ٤٩٣
مصادر المحك ٤٩٤
ملامح الاختبارات مرجعية المحك ٤٩٥
تفسير درجات الاختبار المرجع إلى المحك ٤٩٧
الاختبارات التحصيلية المرجعة إلى المعيار ٤٩٨
مقارنة بين الاختبارات المرجعة إلى المعيار والاختبارات المرجعة إلى المحك. ٤٩٩
مراجع الباب الخامس ٥٠٢
الباب السادس: مقاييس الشخصية ٥٠٥
الفصل الثالث عشر: وسائل التقرير الذاتى ٥٠٦
مقدمة ٥٠٧
اختبار مينسوتا المتعدد الأوجه ٥١٣
اختبارات مشتقة من اختبار الشخصية المتعدد الأوجه ٥٢٠
اختبار مينيسوتا للإرشاد النفسى ٥٢١
نماذج أخرى ممثلة من اختبارات الشخصية المعتمدة ٥٢٣
على وسائل التقرير الذاتى
الفصل الرابع عشر: قياس الشخصية عن طريق ٥٤١
الأحكام والملاحظات المنظمة
مقاييس التقدير وتقريرات القياس الاجتماعى..... ٥٤١
تقديرات الرؤساء والكبار..... ٥٤١
تقديرات الأقران والرسم الاجتماعى ٥٤٧
الملاحظة المنظمة لعينات السلوك ٥٤٩
الاختبارات الموقفية ولعب الأدوار..... ٥٥٣

الفصل الخامس عشر: اختبارات الأداء والأساليب الإسقاطية ٥٥٩
اختبارات الأداء ذات المواقف محددة البنية ٥٦٠
اختبارات المواقف غير محددة البنية ٥٦٤
اختبارات حل المشكلات ٥٦٤
الاختبارات الإدراكية غير محددة البنية ٥٦٦
اختبار بندر جشطالت ٥٦٦
اختبار بقع الحبر لرورشاخ ٥٦٧
اختبار بقع الحبر لهولتيزمان ٥٧٢
اختبار تفهم الموضوع ٥٧٢
اختبار تفهم الموضوع للأطفال ٥٧٦
الاختبارات اللفظية غير محددة البنية ٥٧٧
اختبار تكملة الجمل ٥٧٧
اختبار تداعي الكلمات ٥٧٨
الفصل السادس عشر: قياس الميول ٥٨١
مقاييس الميول ٥٨١
اختبار سترونج للميول المهنية ٥٨٢
اختبار كيودر للميول المهنية ٥٨٨
اختبار لى وثورب ٥٩٢
الميول والشخصية ٥٩٢
مراجع الباب السادس ٥٩٨
خاتمة : تقويم التقويم ٦٠٥
تقويم أدوات التقويم ٦٠٨
نموذج لتقويم الاختبارات ٦١١

البـــاب الأول

الأسس العامـــة

يشمل هذا الباب الفصول التالية

١- ما هو التقويم النفسى

٢- مدخل إلى الإحصاء النفسى

الفصل الأول

ما هو التقويم النفسى ؟

معنى كلمة التقويم :

من الكلمات التى شغلت الباحثين والعامة على حد سواء فى السنوات الأخيرة كلمة «تقويم» ، وهل لاتزال تدل على المعنى المراد منها؟ أم الأصح أن نستخدم كلمة «تقييم» والتى أصبحت أخف على الألسنة وأوسع فى الانتشار؟.

ولعل مصدر الخلاف هو الجذر الثلاثى للكلمة وهو «قوم» ومنه تصاغ كلمة «قيمة» تبعا للقاعدة الصرفية فى اللغة العربية التى ترى أن «الواو إذا وقعت ساكنة بعد حرف مكسور قلبت ياء لتناسب الكسرة» (١٦ أ: ١٨٠ ،١٦ب، ٢١٢)(*)، إلا أن القاعدة العامة المطردة فى الاشتقاق بالنسبة لمثل هذه الكلمات المشتقة هو العودة إلى أصل الحروف فى الثلاثى، وعلى هذا ففى حالة كلمة، قيمة» نعود إلى الأصل «قوم» مرة أخرى ناظرين إلى الواو مرة أخرى فنقول «قوم» ، «وتقويما» .

ولكن لوحظ أن بعض العرب أهملوا النظر إلى أصل الحرف فى الجذر الثلاثى ونظروا إلى حالته الراهنة بعد إبداله فى الكلمة، وهذا ما يسميه بعض النحاة «الاشتقاق على التوهم» ، أو ما يسميه بعض المحدثين «الاشتقاق من المشتق» ، دفعا للغموض وإزالة للبس.

ومن هذه القاعدة الخاصة أجاز مجمع اللغة العربية فى القاهرة أن يقال قيمت الشيء تقييما بمعنى حددت قيمته، وذلك للتفرقة أو إزالة للبس بين هذا المعنى وبين «قومته» بمعنى عدلته وجعلته قويما أو مستقيما.

وقد شاع أيضا فى إصدار الأحكام على الأشياء مصطلح ثالث Assessment وترجمتنا له «تقدير» وقد إرتبط هذا المفهوم بالمجالات الكلينكية حيث يستخدمه الكلينيكيون فى فحوصهم فى حالات المرض أو الإضطرابات النفسية. أو حين يجرى الباحث فى المجال الإكلينيكى دراسة ما.

(*) يشير الرقم الأول إلى رقم المرجع فى القائمة المثبتة فى نهاية كل باب من أبواب الكتاب، بينما يشير الرقم الثانى بعد العلامة (:) إلى رقم الصفحة أو الصفحات فى ذلك المرجع نفسه، وعندما توجد علامة الوصل (،) فإن ما بعدها يشير إلى رقم مرجع آخر وأرقام صفحاته وهذا هو النظام المتبع فى الكتاب كله.

وقد عرف (Sundberg :45)هذا المفهوم بأنه مجموعة من العمليات التى تستخدم بواسطة أخصائيين للوصول إلى تصورات وإنطباعات لإتخاذ قرار نحو سلوك معين وقد يأخذ هذا التقدير صيغة كمية أو كيفية لوصف الظاهرة موضع التقدير (مثل تقديرات المعلمين لطلابهم أو الأقران لأقرانهم أو الرؤساء لمرؤوسيهم) . ويصلون إلى هذه الأحكام من خلال الملاحظة أو إستخدام مقاييس التقدير. وقد يغلب على التقدير الحكم الإنطباعى.

وهكذا توجد فى لغتنا العربية كلمات فصيحة ووجودها يمكننا فى حل مشكلة التداخل فى المعنى والخلط فى الاستخدام فالكلمات الأجنبية فى اللغة الإنجليزية وهى Assessment & Valuation & Evaluation يمكن ترجمتها كما يأتى بالترتيب تقدير تقييم تقويم. فالكلمة الأولى والثانية لا يتجاوز معناها تحديد القيمة أو القدر، كما أنهما ترتبطان إلى حد كبير بذاتية المقدر بينما كلمة تقويم تتضمن معنى التعديل والتحسين والتطوير وتتسم بالموضوعية.

هل هو تقييم نفسى أم تقدير نفسى أم تقويم نفسى ؟

لاشك فى أن علم النفس يتضمن الكثير من عمليات «التقييم» بالمعنى الذى بيناه فى الفقرة السابقة، أو بمعنى إصدار الأحكام على «قيمة» الأشياء أو الأشخاص أو الموضوعات أو الأفكار.

فالإنسان لايكاد يتوقف عن التقييم، وإعطاء قيمة لما يدرك. إلا أن هذا التقييم فى معظمه من النوع الذى يمكن أن نسميه «التقييم المتمركز حول الذات» ، ومعناه أن الشخص يحكم على الأشياء أو الأشخاص أو غير ذلك بقدر ما ترتبط بذاته هو . وقد يستخدم فى أحكامه هذه اعتبارات المنفعة أو الألفة أو الفائدة أو المكانة أو المركز الاجتماعى أو نقصان تهديد الذات أو سهولة الفهم والإدراك، وهذه جميعا تصبغ الأحكام فى مثل هذه الحالة بصبغة ذاتية، وقد يكون بعضها أقرب إلى التخمين لأننا قد نصيب فى أحكامنا هذه وقد نخطئ بنفس الدرجة من الاحتمال.

وقد تكون أحكام الفرد ونتائج تقييمه للأشياء والأشخاص قرارات سريعة لايسبقها فحص وتدقيق كافيين لمختلف جوانب الموضوع أو الشخص الذى يصدر عليه الحكم، وقد تتصف أحيانا بأنها لا شعورية، بمعنى أن الشخص منا وهو يصدر هذه الأحكام لايعى الدلالات والقرائن والأسس التى تعتمد عليها أحكامه. وقد لا يستطيع «التلفظ»

بها مع أنه يكون على درجة كبيرة من الثقة فى صحتها، وقد تكون بالفعل كذلك، وهى بهذه الصفة تكون أقرب إلى الأحكام الحدسية.

وقد يكون التقييم النفسى مرتكزا على أحكام تتوافر لدينا عنها دلالات واضحة فى الذهن وتكون على درجة كبيرة من الوعى بها، معتمدين فى ذلك على فهم واف، وكاف للظاهرة التى نقوم بتقييمها، وعلى تحليل دقيق لعناصرها ومكوناتها، وتكون أحكامنا فى هذه الحالة ذات طبيعة استدلالية(*).

إلا أن الأمر لايتوقف فى بعض الأحيان عن هذا الحد، ففى المجال التربوى - مثلا - تتطلب دراسة السلوك الإنسانى «عمليةإصدار حكم على مدى تحقيق الأهداف التربوية ودراسة الآثار التى تحدثها بعض العوامل والظروف فى تيسير الوصول إلى هذه الأهداف أو تعطيله» (١٤: ٥٤٩) ومعنى ذلك أن إصدار الحكم فى مجال التربية يمكن أن يتبعه إجراء عملى يتعلق بتحسين العملية التربوية، وقد يكون ذلك فى صورة برنامج للتعلم العلاجى مثلا. وهذا كله بالطبع يتجاوز معنى التقييم إلى معنى التقويم.

وبالمثل، حين يجرى الباحث فى المجال الكلينيكى دراسة على الطفل «للحكم» على مستواه العقلى. من حيث التخلف أو التوسط أو التفوق، فإن ذلك يدخل فى باب «التقييم النفسى» لأنه لم يتجاوز مرحلة إصدار الحكم على المستوى وهو من نوع تحديد القيمة أو القدر فقط، وهذا أمر نادر فى الممارسة السيكولوجية المعتادة، لأن الشائع أن يتبع ذلك «إجراء عملى» من نوع ما كإلحاق الطفل بمدرسة للتربية الفكرية أو بفصل للمتفوقين، أو تقديم نوع من التربية التعويضية، أو إعادة تنظيم البيئة التى يعيش فيها وغير ذلك وكل هذا من نوع التقويم النفسى.

ومعنى ذلك أن التقويم النفسى هو الإجراء المعتاد فى المجالات النفسية والتربوية، وهى المجالات التى شاع فيه مفهوم «التقويم» أكثر من غيرها، صحيح أن المفهوم اتسع نطاقه فى السنوات الأخيرة وإمتد إلى مجالات الاقتصاد والاجتماع والسياسة والطب وغيرها إلى الحد الذى دفع بعض المهتمين إلى الكتابة فيما يسمى «علم التقويم Evaluation Science ، وفى جميع الحالات يتجاوز اللفظ حدود «تحديد القيمة» إلى

(*) التمييز بين التخمين والحدس والاستدلال يستند إلى تصور نظرى لأحد مؤلفى هذا الكتاب (راجع فؤاد أبو حطب المرجع رقم ١٢).

التحسين والتطوير والتنمية، ولعل هذا أقرب إلى المعقول، فالباحث فى العلوم الإنسانية - بصفة عامة - لا تتشابه مهمته مع مهمة الناقد الفنى أو الأدبى التى تقتصر على «وزن» العمل الفنى والأدبى «بميزان النقد» ليحدد قيمته وجدواه، ولا تمتد بالطبع إلى اقتراح الطرق التى يمكن بها أن يتحسن العمل الفنى أو الأدبى أو يتطور، وهو عمل أقرب إلى التقييم منه إلى التقويم.

والباحث فى علم النفس ليس «متفرجا» على السلوك الإنسانى ولا تتحدد أهدافه فى مجرد الحكم عليه لتحديد قيمته أو قدره أو مقداره أو تقييمه وإنما تمتد فى كثير من الأحوال إلى تقويم هذا السلوك بالمعنى الذى أشرنا إليه، أى أن التقويم النفسى هو القاعدة العامة والتقييم النفسى هو الاستثناء.

ومن قبيل هذا الاستثناء ما يحدث فى مجال التربية حين تقتصر مهمة المدرسة فيما يتصل بالامتحانات على الحكم على المتعلم بالنجاح أوالفشل.

ومعنى ذلك أن مهمة الأخصائى النفسى أقرب إلى مهمة الطبيب، التى لا تقتصر على مجرد «قراءة» ميزان حرارة الجسم أو مقياس ضغط الدم أو «مشاهدة» الرسم الكهربائى للمخ أو القلب، لمجرد الحكم بالسواء أو المرض، وإنما تتجاوز ذلك إلى التشخيص ثم العلاج.

التقويم النفسى والقياس النفسى:

من المفاهيم التى يحدث بينها خلط كبير فى مجال علم النفس مفهوما التقويم النفسى والقياس النفسى، وقد حددنا فيما سبق معنى التقويم النفسى، فماذا عن القياس النفسى؟

يعرف ننالى (٢٧:٤١) القياس فى العلم عامة وفى علم النفس خاصة بأنه «قواعد استخدام الأعداد بحيث تدل على الأشياء بطريقة تشير إلى كميات من صفة أو خاصية، ومعنى ذلك أن القياس النفسى يعتمد فى جوهره على استخدام الأعداد، إلا أنه فى صورته المحكمة يتضمن فكرة الكم والتى تعنى مقدار ما يوجد فى الشخص من خاصية معينة.

ورغم أهمية هذا التصور الكمى لخصائص الإنسان إلا أنه ليس شائعا فى جميع الأحوال وبالنسبة إلى جميع الخصائص، فبعض الخصائص لايزال يستعصى على التناول الكمى المباشر وتستخدم فى وصفه «لغة كيفية» كاملة وسواء لجأنا إلى الوصف «الكمى» أو الكيفى» للسلوك الإنسانى فإن ذلك يعد من قبيل البيانات التى لابد أن تتوافر للباحث النفسى

قبل أن يصدر أحكامه أو يتخذ قراراته أو يقترح الإجراءات العلاجية والتصحيحية،. وعلى هذا فإن القياس النفسى يوفر للباحث البيانات والمعلومات الكمية التى يعتمد عليها عند القيام بعمليات التقويم التفسى، إلا أنه ليس مصدر المعلومات الوحيد، فإذا لم تتوافر لهذا الباحث الطرق الكمية لوصف السلوك، أو إذا استعصت الظاهرة النفسية على التناول الكمى فلابد له من الاعتماد على الطرق الكيفية فى جمع المعلومات.

ومعنى ذلك أن التقويم أكثر عمومية من القياس، فالتقويم النفسى يعتمد على المقاييس أو على غيرها من طرق جمع المعلومات عن السلوك الإنسانى. وفى كلتا الحالتين يتضمن التقويم إصدار الأحكام القيمية واتخاذ القرارات العملية، كما أن مقادير القياس الكمية أو الصيغ اللفظية الوصفية يمكن استخدامه استخدامات مختلفة تبعا لأهدافنا من عملية التقويم النفسى. فنتائج الاختبار التحصيلى الواحد (والتى قد تتخذ صورة قياس كمى، يمكن استخدامها فى الحكم على مدى تحقيق هدف مثل الوصول بالتلاميذ إلى المستوى المعتاد من الكفاءة، كما يمكن استخدامها فى الحكم على مدى تحقيق هدف مختلف مثل تنمية قدرة معينة، وبالطبع تختلف فى الحالتين خطة الباحث التى يستخدمها للوصول إلى مثل هذه الأحكام وما يتلوها من خطوات عملية.

مستويات القياس:

أشرنا إلى أن القياس يستخدم لغة الكم، إلا أن هذا لايعنى أن جميع المقاييس من فئة واحدة، فقد قام علماء النفس بتحديد أنواع القياس المختلفة ودرجة ملاءمة العمليات الحسابية المعروفة لكل من هذه الأنواع، ولهذا الموضوع أهميته القصوى لأننا لو لم نتناوله ببعض التعمق قد نفترض أن جميع العمليات الحسابية من جمع وطرح وضرب وقسمة يمكن أن تستخدم فى كل نظم القياس. بل قد يستنتج البعض أن القياس يصبح مستحيلا ما لم نستخدم جميع العمليات الحسابية، ولو أخذنا ميدان القياس العقلى منذ بدايته لوجدنا وعيا مبكرا بهذه المسألة فعلى سبيل المثال نجد أنه منذ ظهور الاختبارات المبكرة فى هذا المجال أشار العلماء أنه لا معنى بأن الطفل الذى نسبة ذكائه ٦٥ يكون نشاطه العقلى نصف طفل آخر نسبة ذكائه ١٥٠، فالأطفال يختلفون دون شك كثيرا فى استجابتهم للمواقف التى تتطلب الاستدلال المجرد، إلا أنه لايوجد فى سلوكهم ما يبرر التعبير عن هذه الفروق فى صورة معامل مقدار ٢ أو كسر مقداره ١/٢ أو نسبة مقدارها ٢ : ١، بل لا معنى مطلقا لقسمة نسبة ذكاء على نسبة ذكاء أخرى.

وحينما تنبه علماء النفس إلى هذه الخاصية التى تتوافر فى الأعداد التى يستخدمونها تحققوا من أنه توجد مقاييس مشابهة خارج علم النفس. فالحرارة مثلا تقاس فى العادة بمقاييس مئوية (أو فار نهايتية) فإذا انخفضت الحرارة من ٣٠ درجة مئوية أثناءالنهار إلى ١٥ درجة مئوية أثناء الليل لا نستنتج من ذلك أن الجو أصبح فى منتصف الليل نصفة دفئه أثناء الظهيرة وذلك لأن الصفر فى مقاييس الحرارة صفر اعتباطى ولا يعنى عدم وجود حرارة على الإطلاق، كما أن درجات الحرارة فوق الصفر لايمكن تناولها بنفس الطريقة التى نتناول بها درجات الطول أو الوزن، ومن الواضح أن مقاييس الذكاء أقرب إلى مقاييس الحرارة منها إلى مقاييس الأطوال.

والواقع أن أفضل تصنيف لأنواع القياس المختلفة أو مستوياته وبخاصة فى ميدان علم النفس ذلك اقترحه ستيفن (٤٤) وفيه يقسم الطرق المختلفة لاستخدام الأعداد إلى أربعة أنواع هى المقاييس الاسمية ومقاييس الرتبة والمسافة والنسبة لكل منها قواعده وحدوده وضوابطه، ولكل منها الإجراءات الإحصائية الملائمة له. وقد طور كومبس (٢٩) هذا التصنيف وحدد العلاقات التى تربط بين المستويات المختلفة وأضاف فئات جديدة نعرضها فيما يلى:

١ - **المقاييس الاسمية**: وهذا النوع من المقاييس يقع فى أدنى مستويات القياس فى ضوء مدى ملاءمة الإجراءات الحسابية المعروفة له ومدى تطبيقها فيه. وتسمى المقاييس الاسمية nominal كذلك لأن الأعداد تخصص لتعيين عناوين lables أو فئات categories ينتمى إليها الأشخاص أو الأشياء كأفراد.

ومن أمثلة استخدام الأعداد لتدل على عناوين ما يقوم به الإخصائى الجيولوجى حين يختار عددا من عينات الصخور ويعطيها ١، ٢، ٣... إلخ. وفى هذه الأحوال تستخدم الأعداد كما تستخدم العناوين، ولايتضن بحال من الأحوال أى مطلب من التحليل الرياضى، ومن هذه الاستخدامات أيضا تلك التى تستخدم وتخصص لكل لاعب من فريق كرة القدم، على علامات الطريق أو أرقام العناصر الذرية.

ويجب أن نوكد أن أى مقياس يهتم باستخدام مقصود للأعداد، ومن هذه الاستخدامات المقصود التسمية أو العنونة labling . وفى هذه الأحوال لايوجد أى غرض لاستخدام التحليل الرياضى للأعداد، كما أن الأعداد لا تعد دالة على كميات من خصائص.

ويرتبط باستخدام الأعداد كعناوين استخدامها لتدل على مجموعات أو فئات من الأشخاص أو الأشياء، ومن أمثلة ذلك إعطاء الرمز (١) ليدل على الإناث والرمز (٢) ليدل على الذكور عند التصنيف حسب الجنس، ويمكن أن تستخدم الأعداد فى أى تصنيف آخر للأسوياء والمضطربين، أو حسب المهن، أو حسب الحالة الاجتماعية... إلخ. والفرق الوحيد بين استخدام الأعداد فى التسمية واستخدامهافى التصنيف هى أننا فى حالة التصنيف يتم تجميع أكثر من وحدة بعضها مع بعض ولذلك فإن الوحدات التى تعطى نفس العدد تتشابه فى خاصة من الخصائص، وكما هو الحال بالنسبة لاستخدام الأعداد فى التسمية فإن الأعداد التى تستخدم لتدل على فئات فى نظام معين للتصنيف ليس لها أى مضمون كمى. وفى كلتا الحالتين فإن العملية الحسابية التى يمكن تطبيقها على المقاييس الاسمية هى عملية العد أو التعداد للحالات الفردية، أما الأعداد المستخدمة ذاتها فلايمكن أن تستخدم معها عملية الجمع أو الطرح أو الضرب أو القسمة.

ويرى كومبس أن هذا النوع من المقاييس تحكمه علاقة التساوى، وهذا يعنى أن أى زوج من الأشياء يجب أن ينتمى بوضوح إلى الفئة نفسها أو لا ينتمى إليها، وعادة ما يحكم علاقة التساوى هذه مبدأ الانتظام symmetry بمعنى أنه إذا كانت أ = ب، ∴ ب = أ، ومبدأ التعدى transitive بمعنى أنه إذا كانت أ = ب، ب = جـ، ∴ أ =جـ، وبالطبع تلعب العوامل الثقافية دوراً مهما فى تحديد مثل هذه الفئات الاسمية وابتكار فئات جديدة.

٢ - **مقاييس الترتيب الجزئى:** يرى كومبس (٢٩ : ٤٧٤ - ٤٧٥) أنه فى الفئة الواحدة من فئات المقاييس الاسمية قد نجد ما هو أكثر من مجرد أن الوحدات التى تتألف منها الفئة متساوية فيما بينها ومختلفة عن الوحدات التى تتألف منها فئة أخرى، فقد توجد بعض العلاقات بين بعض وحدات الفئات المختلفة، ومن هذه العلاقات أن وحدات إحدى الفئات قد تكون أكبر (أو أصغر) من وحدات فئات أخرى، فإذا كان لدينا عدد من الفئات المتكافئة وكانت هذه العلاقات موجودة بين كل فئتين منهما فإننا نحصل على مقياس الترتيب الجزئى Partial ordinal.

ولكى نوضح ذلك لنفرض أننا نريد قياس ما يسمى المستوى الاقتصادى - الاجتماعى ، ولنفرض أيضاً على سبيل التبسيط أن هذه الخاصية تتألف من مكونين

هما مستوى الدخل والمستوى التعليمى،. فإننا فى هذه الحالة إذا وجدنا أن الفرد (أ) أعلى دخلا من الفرد (ب)، وأنه فى الوقت نفسه أرقى منه فى المستوى التعليمى يمكننا أن نقول أن (أ ، ب)فى المستوى الاقتصادى والاجتماعى ، وكذلك إذا كان (ب) أعلى فى مكونى الخاصية من فرد ثالث هو (جـ) فإن (ب) لايصبح فقط أعلى من (جـ) فى المستوى الاقتصادى ـ الاجتماعى ولكن (أ) يصبح أيضًا أعلى من (جـ) فيه أى (أ ، جـ)، ومن هذا يتضح أن العلاقة هنا متعدية مثل النوع الأول من المقاييس ولكنها ليست منتظمة asymmetric لأننا فى المثال السابق نقول أنه إذا كان أ أكبر من ب فإن ب ليس أكبر من أ.

لنفرض أيضًا أن لدينا شخصاً رابعا هو (د) مستواه فى مكونى الخاصية أقل من (أ) وأعلى من (جـ)، أننا فى هذه الحالة نستطيع القول أن أ ، د ، جـ . ولكن لنفرض فى الوقت نفسه أن بالرغم من أن (د) أعلى دخلا من (ب) إلا أنه أقل منه تعليما. أننا فى هذه الحالة نواجه مشكلة حقيقية لأننا لانستطيع أن نحدد مباشرة ما إذا كان (ب،د) أو (د ، ب) بالنسبة للمستوى الاقتصادى والاجتماعى، وهكذا يصبح (ب)، (جـ) لايمكن المقارنة بينما ويتخذ مقياس المستوى الاقتصادى الاجتماعى للأفراد الأربعة (أ، ب، جـ، د)صورة الترتيب الجزئى على النحو التالى:

أ

ب د

جـ

حيث تدل على أن الشخص الأعلى له مكانة أكبر من الشخص الأدنى وحيث لايوجد اتصال بين (ب، د) اللذين لايمكن المقارنة بينهما، ويمكن الحصول على هذا النوع من المقاييس حيث تتألف الخاصية أوالسمة من عنصرين أو أكثر لايرتبطان فيما بينهما بعلاقة الإضافة additive أو التعويض Compenastive.

٢ – مقاييس الرتبة: المستوى التالى من مستويات القياس هو ما يسميه ستيفنس مقاييس الرتبة Ordinal scales ، ويستخدم هذا النوع حين يمكن تنظيم الوحدات فى سلسلة تمتد بين الأدنى والأعلى فى الخاصية التى نقيسها، إلا أننا لانستطيع أن نحدد بدقة الفرق بين أى رتبتين. والواقع أن الترتيب هو أكثر صور القياس بدائية (باستثناء

المقاييس الاسمية)، وهذا يعنى أنه أساس لكل الصور الراقية من القياس، ومع ذلك لايحمل إلا القليل من المعلومات عن الخاصية التى نقيسها، فحين يقوم الفاحصون فى موقف المقابلة بترتيب مجموعة من المتقدمين لشغل إحدى الوظائف فإنهم يستخدمون مقياس الرتبة، كما أن المئينيات التى تستخدم كثيرا فى تفسير نتائج الاختبارات هى نوع من الترتيب وتحمل خصائص مقياس الرتبة. وفى هذا النوع من المقاييس لايمكن استخدام العمليات الحسابية المعروفة من جمع وطرح وضرب وقسمة، ولايمكن استخدام متوسط الرتب أو انحرافها المعيارى، وإنما تصلح لهذه المقاييس أساليب إحصائية أكثر ملاءمة، وهى من النوع الذى يسمى الإحصاء اللابارامترى ومن أشهرها معامل ارتباط الرتب.

ويمكن أن يتحول مقياس الترتيب الجزئى إلى مقياس رتبة كامل إذا أمكن مساواة عناصر الخاصية بعضها ببعض. ففى مثال قياس المستوى الاقتصادى ـ الاجتماعى الذى أشرنا إليه يمكننا أن نساوى كل عام من التعليم بجنيه واحد من الدخل الشهرى (١٢ جنيها فى السنة)، مثلا، وفى هذه الحالة نحصل على أساس إجرائى بسيط للمقارنة بين أى شخصين أو تصنيفهم فى فئات متساوية، ويصبح المقياس فى هذه الحالة مقياسا للرتبة مادامت عناصر السمة تخضع للتحويل الخطى البسيط بحيث يحل أحد العناصر محل الآخر على أساس قاعدة أو مبدأ مفهوم.

٤ – **مقياس الرتبة المترى:** من الملاحظ على أنواع المقاييس الثلاثة السابقة أن عناصر المقياس هى فئة من الأشخاص أو الأشياء، وأن العلاقات بينها وبين علاقات التساوى أو «أكبر من»، ولايتضمن مفهوم المسافات بين الفئات، ومعنى ذلك أننا قد نلاحظ أن (أ) أكبر من (ب) وأن (ب) أكبر من (جـ) إلا أننا لا نستطيع أن نحدد ما إذا كانت المسافة بين (أ) و(ب) أكبر أو أقل من المسافة بين (ب) و(جـ)، ولهذا يرى كومبس (٢٩) أن المقاييس السابقة جميعا هى من قبيل المقاييس الاسمية لفقدانها خاصية مهمة فى القياس عامة وهى خاصية تحديد المسافات بين فئات الأشياء أو الأشخاص.

إلا أنه بالنسبة لجميع المقاييس السابقة لو أمكننا تحديد المسافات بين الفئات فإننا نحصل على مستوى أعلى من المقياس، فمثلا إذا كانت العلاقة (أكبر من) تعود على بعض المسافات بين الأشياء المتجاورة فى مقياس رتبة، فإن هذا المقياس يصبح مقياس رتبة مرتبا جزئيا، أما إذا كانت هذه العلاقة تصدق على جميع المسافات

يصبح المقياس مرتبا ترتيبا كاملا ويطلق كومبس (٢٩) على المقياس فى هاتين الحالتين اسم مقياس الرتبة المترى Ordered metric scale ، ويقصد به المقياس الذى يمكن أن نجد فيه ثلاثية من الفئات أ ، ب ، جـ ، د.

والمثال التالى يوضح ذلك: لنفرض أننا نريد قياس سمة «السلطة» كما تتمثل فى سلوك «الرئاسة» فإننا نستطيع ذلك فى ضوء العلاقة «أكبر من» التى أشرنا إليها فى حديثنا عن مقياس الرتبة المعتاد، فإذا كان (أ) يرأس (ب)، (جـ) يرأس الجماعة التى ينتمى إليها (أ) فإننا نقول أن جـ ، أ. وهكذا عن طريق تتبع التسلسل الرئاسى يمكن إعداد مقياس رتبة.

إلا أن مقياس الرتبة البسيط هذا يمكن تحويله إلى مقياس رتبة مترى إذا توافرت لدينا معلومات تحدد لنا الفروق فى مقدار السلطة الذى يمارسه الشخص مقارنا بآخر. ولكى نوضح ذلك نعطى مثالا من الميدان التربوى، لنفرض أن الموجه الفنى يرأس اثنين من المدرسين الأوائل، وأن كلا من هذين المدرسين يرأس ١٠ مدرسين عاديين، فإننا نستنتج فى هذه الحالة أن سلطة الموجه تزيد على سلطة المدرس الأول بمقدار ٢ مدرسين أول، ٢٠ مدرسا عاديا. فإذا انتقلنا فى سلم السلطة من الموجه العادى إلى الموجه الأول الذى يرأس كل منهما أيضًا ١٠ مدرسين عاديين، فإن سلطة الموجه الأول فى هذه الحالة تزيد على سلطة الموجه العادى بمقدار ٤٠ مدرسا عاديا (٣٠ + ١٠)، ٤ مدرسين أوائل، ٣ موجهين عاديين، وهكذا بالنسبة لكل مستوى من مستويات الرئاسة التربوية، ومن هذا المثال يتضح أن الموجه الأول يفوق الموجه العادى فى السلطة بمقدار أكبر مما يفوق به الموجه العادى المدرس الأول، ونستنتج أن الزيادة فى مقدار السلطة من المدرس الأول إلى الموجه العادى أقل من الزيادة من الموجه العادى إلى الموجه الأول.

٥ - **مقاييس المسافة**: يمكن القول أن ما يميز مقياس المسافة interval عن مقياس الرتبة بمستوياته المختلفة التى أشرنا إليها أنه يسمح بتحديد مدى بعد شيئين أو شخصين بعضهما عن بعض فى الخاصية موضع القياس، وأن تكون هذه المسافة متساوية ويحتاج ذلك إلى وضع قواعد معينة يتم الاتفاق عليها لاستخدام الأعداد حتى يمكن تحديد كم الصفة أو الخاصية فى الشىء أو الشخص، ومن أمثلة ذلك أننا يمكننا أن نحصل على مقياس مسافة للأطوال فى جماعة من الأطفال بالنسبة لأقصر طفل

فى المجموعة، وبالتالى فإن أقصر طفل فى هذه المجموعة يحصل على الدرجة صفر، والطفل الأطول منه بخمسة سنتيمترات يحصل على الدرجة ١، والأطول عشرة سنتيمترات يحصل على الدرجة ٢، وهكذا. والإجراء الأقرب إلى الشيوع فى أغلب المقاييس النفسية أن تحدد المسافات فى ضوء بعد كل طفل عن المتوسط الحسابى لأطوال المجموعة بمقادير معيارية كما تتمثل فى الإنحراف المعيارى لنفرض أن المتوسط ١٢٠ سم والانحراف المعيارى ٥ سم، فإننا نستطيع فى هذه الحالة أن نعطى +١ للطفل الذى يزيد طوله عن هذا المتوسط بمقدار ٥ سم، والدرجة -١ للطفل الذى يقل طوله عن المتوسط بمقدار ٥ سم أيضًا. وهذه المسافات حول متوسط الأطوال وغيره من الخصائص يمكن حسابها دون معرفة حقيقية بمدى بعد الأشخاص عن نقطة الصفر.

والواقع أن أغلب المقاييس النفسية والتحصيلية والعقلية من هذا القبيل، فنحن نقارن درجات طالبين فى الاختباروتحدد مدى بعد كل منهما عن المتوسط، إلا أن أهم نواحى القصور فى هذه المقاييس أنها ليس لها صفر مطلق. فقد يحصل التلميذ على الدرجة صفر فى أحد الاختبارات التحصيلية إلا أن ذلك لايعنى أنه لايوجد لديه معلومات على الإطلاق حول موضوع الاختبار، كما أن الطفل الذى أعطيناه صفرا فى الطول فى مثالنا السابق لا يعنى أنه ليس له طول على الإطلاق وإنما الصفر فى هذه الأحوال هو صفر اعتباطى.

ويمكن أن تستخدم مع مقاييس المسافة عمليات الجمع والطرح، إلا أن عملية القسمة بالذات لايجوز استخدامها على الإطلاق، فلا نستطيع أن نقسم الدرجة التى حصل عليها التلميذ (أ) فى الاختبار على الدرجة التى حصل عليها التلميذ (ب) فى نفس الاختبار، لأن القسمة تفترض مقدما وجود الصفر المطلق، وبالتالى لايمكن الحصول على النسبة (كالضعف أو النصف)، ولتوضيح ذلك نضرب المثال التالى:

نفرض أننا طبقنا اختبارا على شخصين فحصل الأول على الدرجة ٨٠ وحصل الثانى على الدرجة ٤٠، ولنفرض أن الباحث الذى أعد هذا الاختبار ضمنه بالصدفة ١٠ مفردات أخرى يسهل على كل من هذين الشخصين الإجابة عليها إجابة صحيحة، ففى هذه الحالة يظل الفرق بين الدرجتين ثابتا، أى ٤٠، فى الحالتين ولكن

معامل الدرجتين لن يكون متساويا، فبدلا من أن يكون ٢ (أى ٨٠÷ ٤٠) فى الحالة الأولى يصبح ١,٨ (أى ٩٠÷٥٠) فى الحالة الثانية، وهكذا فإننا فى الاختبار الواحد لاتوجد لدينا طريقة لإيجاد ما إذا كانت معلومات أحد الأشخاص ضعف معلومات شخص آخر أو نصفها أو ثلاثة أمثالها ، ولكننا حين نفترض أن كل مفردة للمعلومات يتساوى فى جودته مع أى مفردة أخرى، فإننا لا نخرق مبادئ الرياضيات أو المنطق حين نطرح درجة الشخص الأول من درجة الشخص الثانى أو حين نجمع هذه الدرجات ونستخرج متوسطها وانحرافها المعيارى. والواقع أن الطرق المختلفة التى تستخدم فى ترجمة الدرجات الخام فى الاختبارات إلى الأنواع المختلفة من المعايير - وهى الطرق التى سنشير إليها فيما بعد - تعتمد فى جوهرها على عمليتى الجمع والطرح.

٦ - **مقاييس النسبة** : تعد مقاييس النسبة ratio scales أعلى مستويات القياس وفيها تقاس خاصية معينة فى الشىء بوحدة أو مقدار معيارى من هذه الخاصية بهدف معرفة عدد الوحدات المعيارية من هذه الخاصية التى توجد فيه. فمثلا عندما نقيس طول الحجرة بهدف عد الأمتار أو السنتيمترات (وحدات القياس) التى توجد وتتكرر فى هذا الطول، فإننا نستخدم هذا النوع من المقاييس، وتصلح مع مقاييس النسبة كل العمليات الحسابية من جمع وطرح وضرب وقسمة ومع صلاحية هذه العمليات يمكن استخدام الرياضيات العليا كالجبر والهندسة التحليلية والتفاضل والتكامل وكل الطرق الإحصائية المتقنة، وتتوافر فى مقاييس النسبة جميع خصائص مقاييس المسافة بالإضافة إلى الصفر المطلق، ونحن نألف هذا النوع من المقاييس أكثر من غيره لأن جميع الأبعاد الفيزيائية المعروفة كالطول والوزن والحجم يمكن قياسها بهذه الطريقة، ولذلك نستطيع القول أن الشخص الذى طوله ١٨٠ سم ضعف الشخص الذى طوله ٩٠ سم، والواقع أن تسمية هذا النوع من المقاييس باسم مقاييس النسبة جاءت من قابلية هذه المقاييس للقسمة أو التعبير عنها فى صورة نسبة.

وهذا النوع من المقاييس لايوجد إلا قليلا فى علم النفس، ولاتتوافر المقاييس القليلة من هذا النوع إلا حين نقيس الخصائص العقلية أو الوجدانية بوحدات فيزيائية. كأن نقيس زمن الرجع أو التعلم بوحدات زمنية (كالثانية أو أجزاء الثانية).

طرق القياس فى علم النفس:

يهتم القياس فى العلم دائما بجانب معين من جوانب الأشياء، أو خاصية من خصائصها، وعلى وجه التحديد نستطيع القول أننا لا نقيس الأشياء وإنما نقيس خصائصها ويصدق هذا على الإنسان، فنحن لا نقيس الطفل مثلا وإنما نقيس ذكاءه أو اجتماعيته ولهذا التمييز أهميته من ناحيتين: أحدهما أن القياس يتطلب عملية تجريد، فالخاصية تدل على علاقة بين الأشياء وهذه المسألة رغم وضوحها للقارئ المتخصص إلا أنها لاتبدو كذلك لغير المتخصصين، فكثيرا ما نجد خلطا بين خاصية معينة للأشياء وكل الخصائص التى يمكن ملاحظتها فيها، والفشل فى تجريد خاصية معينة يجعل مفاهيم القياس صعبة الفهم، ومن ذلك مثلا يصعب على كثير من الناس أن يفهم أن الجانح والسوى يمكن أن يتوافر لهما المستوى نفسه من الذكاء (كما يقاس باختبارات الذكاء).

والسبب الثانى للتأكيد على أن القياسَ يتناول خاصية معينة هو أن ذلك يجبرنا على ضرورة دراسة طبيعة الخاصية قبل محاولة قياسها لأن من المحتمل ألا توجد هذه الخاصية. فليس من الضرورى أن جميع العبارات التى نستخدمها فى وصف الناس تقابلها خصائص قابلة للقياس.

ويوجد احتمال آخر أن المقياس قد يتناول خليطا من الخصائص لا خاصية واحدة. وهذا هو ما نجده كثيرا فى استخبارات «التوافق» أو قوائم مشكلات الشباب والتى تتضمن مفردات ترتبط بعدد من الخصائص المنفصلة إلا أنه لا تتضمن إلا قليلا بمنطق القياس فى العلم.

والخصائص التى تخضع للقياس فى العلم قد لا تقاس مباشرة بوحدات معيارية منها، فالحرارة تقاس بمقياس مترى، وكذلك الوزن ، كما أن معظم هذه الخصائص من نوع التكوينات الفرضية hypothetical constructs التى تلاحظ هى فى ذاتها مباشرة أى أنها لا تخضع للإدراك الحسى المباشر، وكثيرا من المفاهيم الكبرى فى العلم من هذا القبيل، فالمغناطيسية والنشاط الذرى وانتقال الحرارة والذكاء من نوع التكوينات الفرضية أو التجريدات المعرفية، وبالطبع تحتاج دراستنا لهذه الظواهر وقياسها توافر قدر من الوقائع التى يمكن ملاحظتها أو المؤشرات الخاصة بها، وإذا كان

علماء الفيزياء يعتمدون فى دراستهم لظواهر المغناطيسية والكهرباء والحرارة على مؤشرات مثل تغيير اتجاه إبرة البوصلة أونشاط عداد جايجر أو قراءة الترمومتر، فإن علماء انفس فى دراستهم وقياسهم لتكويناتهم الفرضية كالذكاء، والقدرة الميكانيكية والميل العلمى والاتجاه نحو العمل اليدوى وغيرها يعتمدون على مؤشرات على هذه الخصائص النفسية تنتمى إلى عالم الواقع ويمكن ملاحظتها ملاحظة موضوعية مباشرة هى أساليب الأداء performance والتى تتخذ إحدى صور ثلاث:

(١) الأداء اللغوى كما يتمثل فى النطق والتحدث والتلفظ شفاها وقد يمتد إلى وسائل الاتصال غير اللفظى كالإيماءات والإشارات وغيرها.

(٢) الأداء الحركى كما يتمثل فى نشاط أعضاء الحركة مباشرة كاستخدام الجسم أو الأيدى أو الأصابع أو الأقدام.

(٣) الأداء الفسيولوجى كما يتمثل فى نشاط الأجهزة الجسمية المختلفة كالنشاط الهرمونى أونشاط المخ.

ويرتبط مفهوم القياس بالعد، فإن ما يفعله المرء فى أى مقياس هو عد الوحدات المتشابهة، فيقاس الطول بعدد الوحدات المتساوية من الأمتار (السنتيمترات مثلا)، ويقاس الوزن بعدد الوحدات المتساوية فى صورة جرامات (مقدار جذب الأرض لسنتيمتر مكعب من الماء النقى)، ولاشك أن الخصائص الفيزيائية لبعض أدوات القياس لها فائدتها فى أن تقوم بعملية العد ذاتها بدلا من أن يقوم بها الملاحظ نفسه. وترتبط المقاييس النفسية أيضًا ارتباطاً وثيقًا بالعد، وبالتالى فإن الكرونسكوب يفيد كعداد لأجزاء الثانية فى قياس زمن الرجع، ويعد عد أزواج الكلمات التى يذكرها المفحوص فى تجربة من تعلم الارتباطات الثنائية يصل الفاحص إلى تحديد مقدار ما تم تعلمه، وبالمثل فإننا بعد عد الأسئلة التى أجاب عليها المفحوص فى اختبار معين نصل إلى قياس لقدرته.

ويعد أداء الفردبعد عرض قائمة الارتباطات الثنائية، أو حله للسؤال فى الاختبار مؤشرا على الخاصية النفسية موضع القياس.

ويستخدم علماء النفس عددا من الطرق فى قياس الخصائص النفسية توضح لنا كلها مقدار السمة أو الخاصية فى ضوء مؤشراتها، وأهم هذه الطرق ما يأتى:

١ – تكرار أو احتمال حدوث مؤشر السمة أو الخاصية: يمكن القول أن عدد الاستجابات «الصحيحة» فى اختبار للقدرة مثال لما نسميه تكرار حدوث مؤشر السمة، لأن كل مفردة من مفردات الاختبار هى «فرصة» للدلالة على ما إذا كانت استجابة المفحوص تدفعه إلى الطرف الأعلى من المقياس أوتهبط به إلى طرفه الأدنى أو تضعه فى المستوى المتوسط.

فمثلا إذا أجاب المفحوص إجابة صحيحة على ٦٠ سؤالا من ١٠٠ سؤال فإن قياس السمة بدلالة تكرار حدوث مؤشرها يمكن أن يتحول فى هذه الحالة إلى احتمال فى صورة نسبة مئوية فنقول أن احتمال حدوث مؤشر السمة فى هذه الحالة هو ٦٠ ٪.

فإذا كان العمل الذى يؤديه المفحوص يتألف من وحدات (أسئلة) كثيرة متساوية فى الصعوبة أو السهولة فيمكن أن تقاس السمة بعدد الاستجابات الصحيحة (سواء كانت فى صورة تكرار أو حتمال) التى يؤديها المفحوص فى وقت محدد، وفى هذا نحن نقيس السرعةو وقد تقاس السمة بعدد الاستجابات الخاطئة التى تصدر عن المفحوص فى وقت محدد وفى هذا يصبح الباحث مهتما بقياس الدقة.

٢ – شدة حدوث مؤشر السمة أو الخاصية: وتتمثل طريقة الشدة أو الحدة in-tensity فى اختبارات القدرة فى مستويات صعوبة المفردات التى يستطيع المفحوص الإجابة عليها، وتتمثل فى النشاط العضلى أو الجهد المبذول كما يقاس بسعة الاستجابة magnitude أو ارتفاع وعلو الاستجابة altitude أو قوة الاستجابة strength ، وفى بعض صور النشاط الأخرى فإن المكونات الانفعالية مثل ضغط الدم أو معدل النبض أو توصيل الجلد للتيار الكهربائى أو التوتر العضلى تدل على الشدة. وتدل شدة الاستجابة بالقبول أو الرفض للقضايا الخلافية على قوة الاتجاهات.

وتستخدم هذه الطريقة فى القياس حين يكون على الباحث المبالغة فى تأكيد القوة. ويمكن أن تستخدم كمؤشر للسمة حين يحصل الفاحص على مقياس يتألف من أسئلة مرتبة حسب الصعوبة، ويصبح السؤال فى هذه الحالة ما هو الحد الذى يصل إليه المفحوص ولايتعداه؟ ويمكن إعطاء مثال واضح من ميدان الألعاب الرياضية وخاصية فى قفز الحواجز حيث يرفع الحاجز تدريجيا حتى يصل اللاعب إلى الحد الذى لايستطيع اجتيازه، ويمكن الحصول بهذه الطريقة على مقياس للأداء. ويصدق

هذا على مقياس بينيه للذكاء الذى يتألف من أسئلة متدرجة فى الصعوبة، ويتطلب هذا جهدا كبيرا فى تحديد هذه المستويات، وقد تعطى للأسئلة درجات متناسبة مع صعوبتها بالنسبة للمفحوصين وتصبح درجة المفحوص هى مجموع هذه الدرجات الموزونة للأسئلة التى يجيب عليها إجابة صحيحة أو خاطئة (حسب نظام وزن الدرجات) وقد يعطى للمفحوص درجة رتبة لأصعب أو أسهل سؤال أجاب عليه.

٣ – مدى حدوث مؤشر السمة أو الخاصية: وهذه الطريقة ليست واضحة أو شائعة الاستخدام كالطريقتين السابقتين وتتطلب فى جوهرها تحديد عينة متنوعة من الأسئلة يتألف منها الاختبار ويستخدم فى قياس السمة تنوع الأسئلة التى أجاب عليها المفحوص كما يتحدد مثلا بعدد فئات الاستجابة.

تعريف الاختبارات النفسية ومهامها:

لعل أشهر تعريفين للأختبارات النفسية هما تعريف اناستازى سنة ١٩٩٠ وتعريف كرونباك سنة ١٩٩٠، وأيضا، وما توصل منهما فؤاد أبو حطب لتصور إجرائى شامل لتعريف الاختبار النفسى كما يلى من العرض.

حيث تعرف أنستازى (٢٣-٢٤) الاختبار النفسى أنه «مقياس موضوعى مقنن لعينة من السلوك». ويعرفه كرونباك (٣٠) بأنه طريقة منظمة لمقارنة سلوك شخصين أو أكثر. وقد ناقش فؤاد أبو حطب هذين التعريفين فى موضع آخر (١٢) وتوصل منهما إلى تعريف أكثر شمولا. ونبدأ بعبارة «عينة من السلوك» التى وردت فى تعريف أنستازى فنقول أن الاختبارات النفسية تتشابه مع الاختبارات التى تستخدم فى العلوم الأخرى، بشرط أن تتم الملاحظات على عينة صغيرة جيدة الانتقاء من سلوك الفرد، وفى هذا الصدد فإن المتخصص فى علم النفس يحذو حذو المتخصص ففى علم الكيمياء مثلا حين يختبر دم المريض أو ماء الشرب فيقوم بتحليل عينة أو أكثر منه. فإذا أراد السيكولوجى أن يختبر المحصول اللفظى للطفل، أو قدرة موظف الحسابات على أداء العمليات الحسابية أو مدى التآزر بين العين واليد عند الطيار، فإنه يفحص أداء هؤلاء فى «عينة» ممثلة من الكلمات أو العمليات الحسابية أو الأعمال الحركية. أما مسألة شمول الاختبار السلوك موضوع الاختبار أو عدم شموله فتتوقف على وحدات العينة وطبيعتها، فالاختبار الذى يتكون من خمس مسائل فقط، أو من أسئلة فى

الضرب فقط يعد مقياسا غير جيد للمهارة العددية، واختبار المفردات اللغوية الذى لايتضمن إلا مصطلحات الألعاب الرياضية لا يعطينا تقديرا يعول عليه للمحصول اللغوى للطفل.

وتعتمد القيمة التشخيصية diagnostic والتنبؤية predictive للاختبار على مدى تقنينه وموضوعيته إلا أن الموضوعية والتقنين هما من شروط الاختبارات الجيدة إلى جانب شروط أخرى سنتناولها بالتفصيل فيما بعد وتضمينها فى التعريف يخل بشرط الإيجاز اللازم. إلا أن هذه القيمة التشخيصية والتنبؤية للاختبار تعتمد فى جوهرها على مدى استخدام الاختبار كمؤشر على مجال أوسع وأكثر أهمية من السلوك، فمن النادر أن تكون عينة السلوك التى يشملها الاختبار هدف القياس النفسى فى ذاتها وإنما المهم أن تبرهن على وجود قدر من التطابق الوثيق بين عينة السلوك فى الاختبار والمجال الأكثر شمولا الذى ينتمى إليه (بين معرفة الطفل لقائمة الكلمات فى اختبار الاستعداد اللغوى وإتقانه الكلى للمفردات اللغوية مثلا كما تتمثل فى قراءاته لإحدى مجلات الأطفال).

ويجب أن نلاحظ فى هذا الصدد أن الأمر لايحتاج إلى وجود التشابه التام بين أسئلة الاختبار والسلوك الذى نسعى إلى تشخيصه أو التنبؤ به، وإنما الضرورى هو البرهنة على وجود تطابق تجريبى بينهما (وهذا ما يسمى الصدق)، ولذلك تتفاوت درجة التشابه بين عينة الاختبار وميدان السلوك الذى تنتمى إليه، ففى الأحوال المتطرفة قد يتطابق الاختبار تطابقا تاما مع السلوك الذى يتنبأ به أو يشخصه، ومن ذلك اختبارات التحصيل المدرسى أو الكفاءة المهنية، وتوجد درجة أقل من التشابه تتمثل فى كثير من اختبارات الاستعدادات المهنية، وفى الطرف الآخر نجد الاختبارات التى لا ترتبط ارتباطا مباشرا بميدان السلوك ومنها الاختبارات الإسقاطية، ورغم هذه الاختلافات السطحية فإن جميع هذه الاختبارات تتألف من عينات من السلوك، ومن الواجب إثبات جدواها بالبرهان التجريبى على وجود تطابق بين أداء المفحوص فى الاختبار وغير ذلك من المواقف.

إلا أن الاختبار قد لايكون بالضرورة عينة سلوك فى كل الحالات. ففى ميدان الاختبارات التحصيلية يزداد الاهتمام فى السنوات الأخيرة بما يسمى الاختبارات

المرجعة إلى المحك criterion-referenced tests وتقوم فى جوهرها على مسلمة ملاحظة معظم أو جميع السلوك وليس عينة منه وخاصة حين يكون موضوع الملاحظة وحدة متكاملة من وحدات التعلم المدرسى، وفى هذه الحالة يصنع الاختبار التحصيلى بحيث يشمل ما تم تدريسه فى هذه الوحدة من محتوى وأهداف وليس عينة منهما، وفى ضوء نتائج هذه الاختبارات المرجعة إلى المحك يمكن تشخيص صعوبات التعلم والتأكد من مدى الاتقان، كما يمكن اقتراح الطرق العلاجية لتحقيق أهداف التعلم التى يمكن أن يحرزها أكثر من ٩٠٪ من التلاميذ، وفى هذه الحالة تصبح العملية تقويما وليس تقييما كما هو الحال فى معظم الممارسات التى تعتمد على ما يسمى الاختبارات المرجعة إلى المعيار.

واستخدام كرونباك لعبارة «طريقة منظمة» فى تعريفه للاختبار أكثر دقة من كلمة «مقياس» عند أنستازي، لأن فى قول أنستازى نوع من عدم التمييز بين المصطلح «اختبار» test والمصطلح «مقياس» measure فعلى الرغم من تداخل معانيهما إلا أنهما ليسا مترادفين.

فمن ناحية نجد أن لفظ «مقياس» أكثر عمومية لأنه يستخدم فى كل ميادين البحث السيكولوجى عندما نسعى إلى الحصول على أوصاف «كمية» كما هو الحال فى بحوث الإدراك والإحساس والحكم والمجال السيكوفيزيائى العام، أى أن اللفظ يستخدم فى الأغراض السيكولوجية العامة، بل وفى صميم علم النفس التجريبى، فكثيرا ما نقيس التعلم أو الاستجابة أو المثير وتستخدم فى هذه الأغراض المقاييس الفيزيائية أو النفسية.

ويطلق على المقياس لفظ اختبار فى مجال استخدامه فى ميدان علم النفس الفارق وحده، وعلى هذا فإن مقاييس العتبات الفارقة أو التعلم أو الإدراك يمكن أن تستخدم «كاختبارات» إذا تحول اهتمامنا بها إلى ميدان الفروق الفردية، إلا أن الاختبار يتكون فى العادة من عدد من الأسئلة أو المفردات التى لاتأخذ صورة مقاييس النسبة وإنما قد تكون من نوع مقاييس المسافة أوالرتبة، ومعنى ذلك أنه ليست جميع المقاييس اختبارات إلا عند الاهتمام بعلم النفس الفارق، وفى هذه الحالة فقط يمكن أن يحل لفظ اختبار ومقياس كل منهما محل الاخر.

ومن ناحية أخرى ليست جميع الاختبارات مقاييس، وقد أشرنا إلى معانى القياس ومستوياته وهو فى كل الأحوال يتطلب نوعا من الوصف الكمى سواء كان من نوع الكم المتصل أو الكم المنفصل، وليست جميع الاختبارات من هذا القبيل، فقد نجد بعض الاختبارات التى لا تعطى درجة للمفحوص وإنما يستخدمها الأخصائى النفسى لمساعدته على الوصول إلى وصف لفظى أو كيفى لمفحوص (مثل طرق الملاحظة). ولا يتطلب الأمر فى هذه الأحوال استخدام المقاييس من أى مستوى من المستويات التى ذكرت آنفا.

ومادامت الاختبارات هى فى جوهرها أدوات الدراسة العلمية للفروق الفردية فإنها تسعى إلى المقارنة بين الأفراد كما يقول كرونباك فى تعريفه إلا أن هذه المقارنة لا تكون بين الأفراد فى ضوء معيار norm فحسب، وإنما قد تكون أيضاً داخل الأفراد فى ضوء مستوى standrad أو محك criterion ، كما أن هذه المقارنة لاتكون فى عينة السلوك فقط، كما هو الحال فى الاختبارات المرجعة إلى المعيار، وإنما تشمل أيضاً المقارنة فى «معظم» السلوك أو «كله» كما هو الحال فى الاختبارات المرجعة إلى المحك.

والمعيار هو أساس للحكم على أداء المفحوصين والمقارنة بينهم فى ضوء أدائهم الفعلى، ويأخذ الصيغة الكمية فى أغلب الأحوال، ويتحدد فى ضوء الخصائص الواقعية لهذا الأداء، ومن ذلك استخدام المتوسط الحسابى لدرجات عينة التقنين معيارا لوصف الأداء العادى فى الاختبار، وفى ضوئه تتحدد الأوضاع النسبية للأفراد فنقول أعلى من المتوسط أو أقل من المتوسط أو متوسط.

أما المستوى فيتشابه مع المعيار فى أنه أساس للحكم على الأداء فى ضوء هذا الأداء ذاته، إلا أنه يختلف عنه فى جانبين: أولهما أنه قد يأخذ الصورة الكمية أو الكيفية، وثانيهما أنه يتحدد فى ضوء ما يجب أن يكون عليه الأداء وليس ما هو عليه بالفعل، ومن هذه المستويات ما نجده فى نظم الامتحانات المعتادة حين نقارن درجات التلاميذ فى هذه الامتحانات بنظام النهايات الصغرى والكبرى، أو حين تتحدد تقديرات النجاح قبليا فى صورة ضعيف ومقبول وجيد وممتاز فى ضوء نسب مئوية من النهاية العظمى للمادة الدراسية توضع مقدما ولا تحسب بالطرق الإحصائية

فى ضوء الأداء الفعلى فى الامتحانات، وهذه جميعا وسائل غير دقيقة فى تحديد المستوى، أما أفضل الطرق فتكون حين يقارن الأداء كما يحدده الاختبار بمستوى الجودة أو الاتقان أو التمكن الذى يحدده الهدف التربوى أو التعليمى أو المهنى، ويكون تحديد هذا المستوى فى الأصل قد تم فى ضوء ما يجب أن يكون عليه الأداء.

أما المحك فهو أساس خارجى مستقل للحكم على الأداء فى الاختبار، وقد تكون هذه المحكات كمية أو كيفية، فمثلا لكى نحكم على نجاح برنامج تعليمى أو تدريبى فى تحقيق أهدافه يمكن مقارنة أداء التلاميذ فى الاختبارات التحصيلية المرتبطة بهذا البرنامج بمستويات الكفاية الإنتاجية التى تتحدد فى الميدان الفعلى للعمل.

وهكذا توصل فؤاد أبو حطب (١٢) إلى تعريف يقترحه لمفهوم الاختبار وفيه يقول:

«الاختبار النفسى هو طريقة منظمة للمقارنة بين الأفراد أو داخل الفرد الواحد فى السلوك أو فى عينة منه فى ضوء معيار أو مستوى أو محك».

تصنيف الاختبارات والمهام النفسية:

يمكن أن نقسم الاختبارات على أسس كثيرة منها الشكل أو الغرض أو المحتوى ويقسمها كرونباك قسمين:

١ – اختبارات الأداء الأقصى maximum performance التى تستخدم إذا كنا نريد أن نعرف إلى أي حد يستطيع الشخص أن يقوم بأداء ما إلى أقصى قدرته، ويمكن الإشارة إليها باسم اختبارات القدرة.

٢ – أما القسم الثانى فيتضمن الاختبارات التى تهدف إلى تحديد الأداء المميز typical performance وهى تقيس ما يحتمل أن يفعله الشخص فى موقف معين أو فى نوع معين من المواقف، وتدخل فى هذا النطاق اختبارات الشخصية والميول.

١ – اختبارات الأداء الأقصى (أو اختبارات القدرة):

أهم ما يميز هذه الاختبارات أن الشخص يشجع فيها ليحصل على أحسن درجة يمكنه الحصول عليها، وتشمل اختبارات الذكاء واختبارات القدرات المختلفة سواء كانت عقلية أو حركية، ومن هذه الاختبارات مجموعة كبيرة تسمى مقاييس القدرة

العقلية العامة وهى تهدف إلى قياس تلك القدرات العقلية العامة فى معظم أنماط التفكير والتعلم (وهى بهذا المعنى اختبارات ذكاء)، وهناك اختبارات القدرات الأكثر تخصصاً وهى القدرات ذات القيمة فى مهام محدودة، منها مثلا الفهم الميكانيكى، والإحساس بالدرجة الصوتية والمهارة الأصبعية، وهى تسمى اختبارات القدرات المتخصصة، وهذه الاختبارات قد تستخدم منفردة أحيانا كما قد تستخدم مجتمعة، وقد تستخدم مجموعة منها تقيس قدرات خاصة متعددة لتساعد فى رسم بروفيل لقدرات الفرد.

أما اختبارات الكفاءة proficiency فتقيس القدرة على أداء مهمة أو عمل ما له أهميته فى ذاته: مثل قراءة الفرنسية، أو العزف على البيانو أو إصلاح خلل فى ماكينة طائرة، وحيث أن هذه الاختبارات تقيس أداء أشخاص تدربوا على العمل المقاس فإنها تسمى أحيانا اختبارات التحصيل achievement.

وهناك اختبارات الاستعدادات aptitude وتستعمل هذه الاختبارات للتنبؤ بالنجاح فى مهنة أو تدريب معين، فهناك مقاييس للاستعداد الهندسى مثلا، أو الاستعداد الموسيقى أو الاستعداد للجبر، وهكذا. ولا تختلف هذه الاختبارات من حيث الشكل عن الأنواع الأخرى، فاختبار الاستعداد الهندسى مثلا قد يتضمن أقساما تقيس القدرة العقلية العامة، وأخرى تقيس الفهم الميكانيكى أو المكان وأقساما تقيس التحصيل فى الرياضيات، ويسمى الاختبار على أى حال، اختبارا تحصيليا إذا كان يقيس نجاح الشخص فى الدراسة السابقة، ويسمى اختبار استعداد عندما يستعمل للتنبؤ بنجاح فى عمل المستقبل.

٢ - اختبارات الأداء المميز:

لا تستعمل اختبارات الأداء المميز لمعرفة ماذا يستطيع الشخص أن يفعله بل لمعرفة ما يعمله بالفعل وطريقة عمله (مثلا فى الاختبار لإدارة عمل ليس المهم ما يمكن أن يعمله الشخص المرشح ولكن ما يعمله بالفعل وأسلوبه فيه). وهناك بعض الخصائص الشخصية، مثل الصبر أو الأمانة أو سعة الأفق لا تستطيع اختبارات القدرة أن تعرفنا بها، على الرغم من أهميتها، فالأداء المميز أو السلوك الفعلى هام من غير شك إذا نحن أردنا أن نقدر الاستعداد للنجاح فى عمل ما.

وإذا كانت الدرجة العالية شيئا مرغوبا فيه فى اختبارات القدرة، ففى اختبارات الأداء المميز لا نستطيع أن تقول أن استجابة ما فى اختبار من اختباراتها «جيدة» إذ ليس هناك ما هو «جيد» أو «ردىء» مثلا فى أن يكون عند شخص ما ميل للهندسة، وكذلك نجد أن بعض الناس يغلب عليهم سلوك السيطرة (وهذه من صفات الشخصية)، ونحن لا نستطيع أن نحدد درجة معينة من السيطرة هى الأنسب والأفضل لأن العالم يتسع لأناس مختلفى الأنماط والصفات.

ونستطيع القول بصفة عامة أن سلوك الشخص المميز له هو أفضل مفتاح لشخصيته ، فالعادات السلوكية لها قيمة تنبؤية فى ذاتها، فما فعله الشخص من قبل يكون أميل إلى أن يعيده ويكرره فى مواقف تالية فيما بعد، ولكن يعترض معظم علماء النفس على افتراض أن عادات الشخص المنظورة هى شخصيته، ذلك أن مواقف جديدة تنشأ دائما، وأن أى وصف للعادات المألوفة عند شخص ما لا تدل بصورة مباشرة على ما سيفعله فى موقف جديد، فربما عرف عن شاب أنه يكره البنات ثم تأتى بنت بعينها فتأسره بحبها ويتعلق بها، والمعالج النفسى الذى يستطيع أن ينشئ علاقة متينة وعميقة وينجح مع كثير من عملائه قد لا يستطيع أن ينشئ مثل هذه العلاقة مع شخص معين لأنه يثير عنده - أى عند المعالج - الكره والغضب مثلا، ولكن مع هذا، ينبغى أن لا تنسينا هذه الاستثناءات أو الظروف الشاذة أن السلوك وردود الأفعال للمواقف المختلفة تعتبر حقا انعكاسات لتركيب شخصية أساسى ومتسق، وهذا التركيب الأساسى والمتسق ينبغى استنتاجه من السلوك، ويأمل العالم النفسى أنه عندما يفهم تركيب الشخصية هذا سوف يستطيع التنبؤ باستجابات الشخص وسلوكه حتى فى المواقف الجديدة.

ويعتبر قياس الأداء المميز أمرا شاقا إلى درجة بعيدة، وقد حقق النجاح بنسب متفاوتة، وبطرق متعددة، وهذه الطرق يمكن تقسيمها إلى قسمين. الأول: ملاحظات السلوك، والثانى: وسائل التقرير الذاتى، أما ملاحظات السلوك فهى محاولات لدراسة الشخص وهو فى حالة سلوكه الطبيعى، والملاحظات يمكن أن تتم فى موقف اختبار مقنن أو فى موقف غير مقنن أو طبيعى. أما الملاحظة المقننة فتتطلب أن يوضع كل شخص فى الموقف نفسه الذى يوضح فيه الآخرون، فيمكن أن تلاحظ «الشخصية» فى موقف اختبار عقلى، أو أثناء مناقشة جماعية، كما أنه من الممكن أن تبتكر مهام أو

أعمال تتيح مجالا طيبا للملاحظة، وتسمى هذه المهام عندئذ اختبارات الأداء للشخصية.

والملاحظة المقننة تسمح بمقارنة شبه تامة لأشخاص لايمكن أن يلاحظوا عادة فى ظروف متشابهة، هذا فضلا عن أنها تكشف عن خصائص لا نشاهدها إلا عرضا فى حياتنا اليومية، وقد استعملت هذه الطريقة مثلا فى ملاحظة ردود الفعل المميزة للحبوط وذلك بأن يتاح للشخص أن يبدأ عملا ولكن يحال بينه وبين أن يحقق غرضه أو يصل إلى هدفه بشكل ما ويعطى رد فعله أوالطريقة التى يسلك بها فى موقف الحبوط ذلك صورة واستبصارا فى مدى ضبطه الإنفعالى.

وإذا كان للملاحظة أن تظهر السلوك المميز فعلا فلابد أن لا يعلم الشخص الخاصية أو الصفة التى يراد ملاحظتها أوتجرى ملاحظتها فعلا، ففى موقف الحبوط المشار إليه منذ قليل من الممكن أن نذكر للشخص أو الأشخاص الذين نختبر ضبطهم الانفعالى، إننا نقيس ذكاءهم، على أننا ينبغى أن نؤكد عند هذه النقطة أن هذا لايجوز إلا فى مجال البحث العلمى والتجريب، ولابد أن يذكر للأشخاص الذين تجرى عليهم التجربة الهدف وما تم فيها بعد أن تنتهى التجربة وكذلك يجوز استخدام هذه الطريقة فى مجال الإختيار لعمل ما لمعرفة مدى توفر صفة أو صفات معينة مناسبة لعمل معين.

وكذلك يمكننا أن نحصل على معلومات عن السلوك المميز بملاحظة عينات من سلوك الشخص العادية فى نشاطاته اليومية فى الميدان كما هى، كأن نلاحظ الأطفال فى الملعب حيث تظهر عاداتهم وشخصياتهم، وكذلك الموظفين فى مكاتبهم، أو العمال فى مصانعهم، وهكذا. وطريقة الملاحظة الميدانية هذه يمكن أن تستخدم فيها أدوات تسجيل مقننة (وأعلى هذه الأدوات من حيث تقنينها التصوير السينمائى)، إذ يستخدم فيها التسجيل والحكم الانطباعى.

أما القسم الثانى من طرق قياس الأداء المميز فهو وسائل التقرير الذاتى والأساس الذى تقوم عليه هذه الوسائل هو أنه لما كان الشخص تتاح له فرصة كبيرة ليلاحظ سلوكه الذاتى، فإنه يستطيع إذا أراد، أن يعطى تقريرا له قيمته عن سلوكه المميز ويساعد فعلا فى التعريف بهذه الناحية، وتستخدم الاستخبارات أو الاستبيانات ques-

tionnaires كوسيلة من وسائل التقرير الذاتى عن الأداء المميز، على أن المشكلة الرئيسية فى التقرير الذاتى، إذا كنا سوف نفسره على أنه صورة للسلوك المميز، هى الأمانة، فإذا حاول الشخص أن يعطى أفضل صورة ممكنة عن نفسه بدلا من وصفها كما هى فإن الاختبار يفشل فى تحقيق غرضه، وحتى عندما يحاول أن يكون صادقا فلن يخلو تقريره عن ذاته من خطأ تحريف أو تحيز.

ومعظم استخبارات التقرير الذاتى تعطى صورة تكاد تكون كاملة عن الشخصية، وبعضها متخصص فى موضوعه، فهناك مثلا استخبارات قياس عادات الدراسة، واستخبارات الميول، واستخبارات الاتجاهات الاجتماعية، وهكذا، وهناك نوع آخر من استخبارات التقرير الذاتى تسمى استخبارات التكيف أو اختبارات الخلق، وهذه الأسماء تدل على الطريقة التى تفسر بها درجات الاختبار ولكنها لا تميز شكلا معينا للاختبار.

ومن المتفق عليه بصفة عامة أن تستعمل استخبارات الشخصية كلمة اختبار فى عناوينها، وذلك منعا لسوء استعمال غير المدربين على إعطاء هذه الاختبارات وتفسيرها وسوء فهم لوظيفة الإستخبارات فى التقويم النفسى.

وسواء كانت الإختبارات المستخدمة من نوع الأداء الأقصي أو الأداء المميز فيمكن تصنيفها أيضا من حيث الشكل والمحتوى وطريقة إستجابة المفحوص عليها ومن حيث الكيف والعمليات المتضمنة والوظيفة النفسية Function

١- من حيث الشكل Form أى طريقة عرض وإعطاء الاختبارات أو المقاييس. وهنا نجد التصنيف الأساسى إلى الاختبارات الفردية والإختبارات الجماعية، ويعد الاختبار الفردى نوع من المقابلة Interview حيث يقوم الفاحص بتوجيه الاسئلة للمفحوص وتسجيل إستجاباته وتقديرها. أما الاختبار الجماعى فيمكن تطبيقه على عدد كبير من المفحوصين فى نفس الوقت. وتصلح الاختبارات الجماعية للتطبيق الفردى ولكن العكس غير صحيح حيث لا تصلح الإختبارات الفردية للتطبيق الجماعى.

٢- من حيث المحتوى Content أى المادة التى تصاغ منها مفردات الإختبار. وهنا يتضح التمييز الأساسى بين الاختبارات اللغوية والإختبارات غير اللغوية . ويجب أن نلاحظ هنا أن هذا التصنيف ليس مطابقا للتصنيف المرتبط بالأداء فى التصنيف الخاص بالأداء فى الاختبار، فإختبارات الورقة والقلم قد تكون لفظية أو غير لفظية،

وكذلك أيضا فى الإختبارات العملية وعادة ما تتكون مادة اختبارات الورقة والقلم غير اللغوية من الصور أو الرسوم وتتخذ تعليماتها صورة الإيحاءات أو الإشارات، أما الإختبارات العملية اللغوية فمن أشهر امثلتها القراءة الجهرية. كذلك يمكننا أن نميز داخل الفئات الاساسية للمحتوى فئات أخرى مثل الاختبارات اللفظية فى مقابل الاختبارات العددية، وإختبارات الصور فى مقابل اختبارات الرسوم والأشكال الهندسية.

٣- من حيث الأداء Performanceأى طريقة الإستجابة التى تصدر عن المفحوص وهنا نميز بين اختبارات الورقة والقلم (الكتابة) والإختبارات العملية. وفى النوع الأول يفكر المفحوص فى المشكلات التى تعرض عليه تفكيرا ضمنيا أو مضمرا ثم يسجل نتائج تفكيره، أما النوع الثانى فيقوم المفحوص بمعالجة المواد التى يتألف منها الإختبار معالجة صريحة (مثل أن يطلب من الطفل بناء برج من أربع مكعبات فى مقياس بينيه).

٤- من حيث الكيف: Quality

وهنا نميز بين إختبار السرعة وإختبارات القوة (ويمكن أيضا إستخدام الإختبارات الموقوتة فى مقابل الإختبارات غير الموقوتة). وتعتمد درجة المفحوص فى اختبارات السرعة على عدد من الاسئلة التى يستطيع المفحوص الإجابة عليها بطريقة صحيحة فى الزمن المسموح به. أما إختبارات القوة فإن درجة المفحوص تعتمد على قدرة المفحوص على الإجابة على الاسئلة الصعبة بدقة. وبالطبع توجد بعض الاختبارات التى تؤكد على سرعة الإستجابة ودقتها فى حل الأسئلة الصعبة.

٥- من حيث العمليات والوظائف النفسية Functions

ترتبط الوظائف والعمليات بكل من الأداء الأقصى والأداء المميز وبالطبع تختلف القدرات والسمات المختلفة تبعا للنظريات المختلفة فى كل مجال، كما ستختلف الاختبارات المعده لكل قدرة أو لكل سمه فى مجال الشخصية. إلا أنه فى كل إختبار لابد لنا من التمييز فيها بين الفئات الأربع السابقة والتى تتوافق فيما بينها بحيث يدل الإختبار الواحد على شكل ومحتوى وأداء وكيف فى وقت واحد. وبذلك يصبح عدد الاختبارات المتوقعة لقياس قدرة معينة أو سمه معينة ١٦ إختبارا (٢×٢×٢×٢) (١٢: ٥٩-٦١).

ويوضح الجدول التالى هذا التصنيف حيث تدل كل خانة فى هذا الجدول على إختبار معين أو مهمة معينة لهذه القدرة، فالخانة (١) تدل على اختبار ورقة وقلم من النوع الفردى ذى المحتوى اللغوى الذى يعتمد على السرعة فى حل مفرداته وهكذا.

(جدول ١)

تصنيف مقترح لإختبار أو مهام إحدى القدرات أو السمات

الوظيفة	الشكل	المحتوى	الأداء			
			ورقة وقلم (كتابى)		عملى	
			سرعة	قوة	سرعة	قوة
أ- أداء أقصى	فردي	لغوى	١	٢	٣	٤
		غير لغوى	٥	٦	٧	٨
	جماعى	لغوى	٩	١٠	١١	١٢
		غير لغوى	١٣	١٤	١٥	١٦
ب- أداء مميز	فردى	لغوى	١	٢	٣	٤
		غير لغوى	٥	٦	٧	٨
	جماعى	لغوى	٩	١٠	١١	١٢
		غير لغوى	١٣	١٤	١٥	١٦

أ: جدول الأداء الأقصى

ب: جدول الأداء المميز

عن فؤاد أبو حطب ١٩٩٦ مع بعض التعديل

ويصنف جنسن (٣٥أ : ٣٧٣-٣٧٤) الإختبارات النفسية من حيث التأثير الثقافى فى أداء الإختبارات النفسية إلى فئتين هما الإختبارات المتحررة من الثقافة Culture Free فى مقابل الإختبارات المشبعة بالثقافة Culture Loaded ويتسم النوع الأول بالعمق (altitude) ويقيس القوة العقلية والاستدلال وحل المشكلات والقدرة على التلخيص، بينما النوع الثانى المشبع بالثقافة فيتم بالإتساع، وكثرة المعلومات المقاسة والسهولة وعدم التعقيد. ويتطلب هذا التصنيف دراسات لإثبات استقلال هذين النوعين عامليا.

البرنامج التقويمى:

البرنامج التقويمى هو خطة منظمة للحصول على المعلومات والبيانات الصحيحة عن الأفراد موضوع التقويم، ولكى نوضح طبيعة البرنامج التقويمى نعطى مثالا من المجال التربوى، حيث يعنى البرنامج التقويمى مجموعة أدوات القياس والتقويم التربوية والنفسية الشاملة والمتنوعة التى تطبقها مجموعة من المختبرين والمرشدين النفسيين والتربويين والمشرفين الاجتماعيين والمدرسين، وفق نظام مرسوم، يحقق حصول المؤسسة التربوية على معلومات وبيانات ترتب وتنظم بحيث يستطيع المشرفون على العملية التربوية أن يستعملوه فى اتخاذ أنسب القرارات للمتعلمين وللمؤسسة التربوية نفسها، أى أن البرنامج التقويمى يشمل: أدوات قياس، ومختبرين، وخطة للاختبار والتقويم، وبيانات منظمة، وطرق الاستفادة من هذه البيانات، ثم علاقات اجتماعية وتفاعل اجتماعى بين هيئة التقويم وبين سائر أعضاء المؤسسة التربوية، وبينها وبين مؤسسة اجتماعية أخرى فى المجتمع تقتضى طبيعتها وطبيعة العملية التقويمية الاتصال بها.

ويجب أن تكون نظرتنا إلى البرنامج التقويمى نظرة صحيحة فيكون العمل فيه محققا لأهدافه وأغراضه، والنظرة الصحيحة إلى البرنامج التقويمى هى التى تعتبره متكاملا مع النظام التربوى كله، كما تعتبره جزءا من التعليم ومن التخطيط التربوى، أى أنه لا يصح أن ننظر إلى البرنامج التقويمى على أنه شىء منفصل عن تيار العمل التربوى فى مؤسساته المختلفة أو منفصل عن عملية التعلم، أو خارجها أو فوقها، ولن يتحقق نجاح لبرنامج تقويمى تربوى لا يقوم على هذا الأساس الكبير: التكامل مع النظام التربوى ومع عملية التعليم ذاتها.

والمسئول عن وضع برناج تقويمى وتسييره، فى أى مستوى من مستويات العملية التربوية، فى مدرسة أو معهد أو كلية أو منطقة تعليمية أو قطر بأسره، يمكن أن يسير وفق خطوات عامة فى بنائه للبرنامج التقويمى على أن تهديه فى خطواته نظره صحيحة إلى التقويم. أما هذه الخطوات فيمكن تلخيصها فيما يلى:

١ - **تحديد الأهداف**: على هذا المسئول أن يبدأ بتحديد أهداف العملية التربوية العامة التى يضعها المجتمع ويتفق عليها وإلتى تحددها السلطات التربوية المسئولة التى تعبر عن أهداف هذا المجتمع وحاجاته، ثم عليه أن يحدد الأهداف التربوية الخاصة بالمؤسسة التى سيتولى إقامة برنامج تقويمى فيها. وبعد أن يحدد هذه الأهداف العامة والخاصة عليه أن يقوم بتعريف القائمين على العمل التربوى فى المؤسسة بهذه الأهداف، ويشاركهم الرأى والمناقشة فى مبرراتها وتفاصيلها ووسائل تحقيقها، ولابأس بأن يوضع دليل مطبوع عن البرنامج التقويمى وعلاقته بهذه الأهداف يفيد منه المدرسون والآباء ومن يعنيهم الأمر.

٢ - **جمع البيانات**: وبعد تحديد الأهداف والتعريف بها، عل المسئول عن التقويم أن يبدأ عملية اختيار الاختبارات التى يمكن الحصول عليها والتى تناسب كل هدف من هذه الأهداف ، مراعيا تنوع هذه الاختبارات وتمثيلها لكل الجوانب التى يراد قياسها، ومن الضرورى مراعاة الشروط الأساسية فى اختيار الاختبارات والتى سنتناولها بالتفصيل فيما بعد.

أما إذا لم تكن هناك اختبارات أو مقاييس مناسبة لأهداف معينة، فعلى القائم على البرنامج التقويمى بالاشتراك مع هيئة مساعدة له من المتخصصين فى القياس، إن وجدت أو من المهتمين فى المؤسسة التربوية، أن يبدأ فى وضع اختبارات ووسائل قياس خاصة ومناسبة، ويمكن بطبيعة الحال الاستعانة ببعض العلماء المتخصصين فى القياس النفسى فى وضع هذه المقاييس، وأن يراعى فى جميع الأحوال شروط بناء هذه الأدوات.

والهدف من اختيار الاختبارات أو بنائها هو استخدامها فى جمع البيانات اللازمة للتقويم، وقد يحتاج الأمر إلى بيانات أخرى لاتوفرها الاختبارات مثل الوثائق والسجلات المدرسية ونتائج الامتحانات السابقة، وهذه البيانات وغيرها لابد من توافرها حتى يكون الحكم على مدى تحقيق الأهداف التربوية سليما.

وعند إعطاء الاختبارات وتطبيقها لابد من انتقاء العينات موضع البحث انتقاء سليما، كما لابد أن تراعى الشروط الخاصة بإعطاء الاختبارات وبهذا تبدأ عملية الحصول على المعلومات والبيانات عن الجوانب التربوية المختلفة موضع التقويم.

وبعد الحصول على هذه البيانات من مصادرها الصحيحة لابد من تنظيمها بطريقة تجعل الاستعانة بها والاستفادة منها مسألة يسيرة تساعد فى اتخاذ القرارات الملائمة له.

وقد يحتاج الأمر إلى تنظيم برنامج للتدريب الدورى لمن يقومون بإعطاء الاختبارات على الأساليب الجديدة، التى يمكن الاستفادة منها فى البرنامج التقويمى، ويدخل فى هذا تدريب المدرسين الذين يبدون اهتماما خاصاً على استعمال وتطبيق بعض الاختبارات والطرق المستحدثة.

٣ - **إصدار الأحكام**: فى ضوء المقارنة بين البيانات التى يحصل عليها االمسئول من التقويم باستخدام الأدوات الملائمة والأهداف التربوية المحددة يتم إصدار الأحكام، وبالطبع فإن هذه الأحكام قد لا تكون مهمة خبير التقويم منفردا، وإنما قد يشاركه فيها عدد من المسئولين فى المنظومة التعليمية، ومنهم المتعلمون أنفسهم، وتزداد هذه المقارنة يسرا إذا كان الهدف قد صيغ فى عبارات إجرائية، لأن هذه الصيغة تتضمن تحديدا لمستوى الجودة والإتقان أو التمكن، بالإضافة إلى أنها تسهل لنا اختيار الأدوات التى ترتبط ارتباطا مباشرا بالهدف.

٤ - **إجراءات التطوير**: تعد خطة التطوير والتنمية والتحسين والعلاج فى ضوء الأحكام التى تم الوصول إليها، وقد تشمل قرارات التطوير تقديم برنامج فى التربية التعويضية أو التعلم العلاجى أو إيداعا فى إحدى المؤسسات أو التدخل بأساليب الإرشاد النفسى أو العلاج النفسى المختلفة.

خصائص البرنامج التقويمى:

هناك سمات وخصائص لابد أن تتوفر فى أى برنامج تقويمى تربوى حتى يحقق الأهداف المرجوة منه ليؤدى وظيفته بنجاح (٨ : ٣٨ - ٤٦).

وأهم هذه الخصائص:

١ - **الشمول**: يجب أن يراعى البرنامج التقويمى الشمول، وبيان هذا أنه لا ينبغى أن ينحصر الاهتمام فى القياس على المعارف والحقائق والمفاهيم، بل يتسع

ليشمل الاتجاهات والميول والتفكير الناقد والتوافق الشخصى والاجتماعية، وأن تستعمل شتى الأساليب بتوازن، والحق أنه إذا كان اهتمام المشرف على عملية التعليم يهتم بالتلميذ اهتماما كليا شاملا، فمن اللازم أن يكون اهتمام البرنامج التقويمى أيضاً شاملا لكل جوانب التلميذ وأنشطته.

٢ - **التنظيم** : يجب أن تجمع نتائج الاختبارات وبياناته وتنظم وتحلل بحيث يصبح تفسيرها ممكنا وذا مغزى، فنتائج البرنامج التقويمى سواء كانت كمية ، أى فى صورة درجات أومذكرات صغيرة أو ملاحظات عابرة، أو كيفية، يجب أن تلخص فى إطار واضح من التقديرات وأن تتحول إلى صورة إحصائية أو رسوم بيانية أو تقارير لغوية، وأن توصف اتجاهات هذه النتائج باستخدام الطرق الملائمة من التحليل الإحصائى أو تحليل المحتوى، وأن تقدم فى جميع الأحوال فى صورة قابلة للفهم والنقل والاتصال بحيث تعطى صورة عن الفرد يمكن مقارنتها بصورة سابقة له، فيساعد هذا على التعرف على اتجاهات نموه فى المجالات المختلفة، وكذا مقارنتها بصور زملائه الآخرين.

٣ - **التكامل** : لابد أن تتكامل المعلومات فى استعمالها عند القيام بحكم معين، وعندما نضع خطة عمل بناء على هذا الحكم، فالمعلومات الخاصة بالنواحى الصحية والتوافق العاطفى والاجتماعى والاتجاهات وغيرها يجب أن تربط مثلا بنتائج اختبارات التحصيل أو اختبار تحصيل بعينه، وكل ما لدينا من بيانات ومعلومات عن الفرد لابد أن نعامله باعتباره وحده يوضح ويكمل بعضها بعضا.

٤- **الاستمرار**: يتميز البرنامج التقويمى الناجح بالاستمرار، أى أنه فى المؤسسة التربوية الحديثة يعتبر عملية مستمرة، فالملاحظات اليومية والتقديرات والاختبارات المستمرة هى التى تشكل العمليات التقويمية، وعن طريقها يحاول المدرس أن يقوم نمو التلاميذ ويوجهه. فلا ينبغى أن تكون هناك فترة محددة للتقويم ثم يتوقف بعدها أو ينتهى تقويم التحصيل المدرسى مثلا فى يوم حساب معلوم أو فى أيام معدودة يوضح فيها الميزان، والحكم الذى ينطق به لايفيد المتعلم بأن يدله على مواطن الضعف والقوة، أو الصحة أو العلة فى أسلوب التحصيل والمراجعة، فلا يفيد هذا الحكم متعلم أو معلم وإنما هو قول فصل، أما النهج القويم فى تقويم التحصيل خاصة، فهو أن يتداخل مع عملية التعلم نفسها، يسبقها ، ثم اثناءها خطوة خطوة حيث يجدى القياس والمقارنة

ويكون حقا قولنا أن التقويم السليم ضرورى للتعلم السليم، بل وضرورى لفاعلية المؤسسات التربوية على اختلاف أنواعها ومستوياتها، ويتمثل ذلك خاصة فى الاهتمامات الحديثة بما يسمى التقويم التكوينى.

٥- **الوظيفية**: لابد أن يعرف كل أعضاء المؤسسة التربوية التى يعمل فيها البرنامج التقويمى بما يمكن أن يقدمه لهم هذا البرنامج من خدمات، فيعمل مثلا على أن يعرف التلاميذ والمعلمون، بل وأولياء الأمور، أنه من الممكن إعطاء اختبارات خاصة غير تلك التى تطبق على كل التلاميذ، أو يمكن إعطاء اختبارات إضافية إذا أراد التلميذ أن يتعرف على بعض جوانب شخصيته، وربما كان أخذ مثل هذه الاختبارات بابا يدخل منه التلميذ إلى مناقشة بعض مشكلاته فيما يتصل بعلاقته بزملائه أو بمدرسيه أو بعمله المدرسى أو حياته الأسرية.

البرنامج التقويمى إذن دعامة رئيسية من دعائم العملية التربوية ويعتمد نجاحها إلى حد كبير على نجاحه، لهذا كان من الواجب على المسئولين عن العملية التربوية أن يعطو البرامج التقويمية التربوية كل جهد أو مال تستحقه، وهو جهد ومال مبذول فى خير وجه، وأنه لمعط ثماره أضعافا مضاعفة.

أنواع التقويم(*)

ينقسم التقويم طبقا لأهدافه إلى أربعة أنواع هى:

التقويم فى منظومة التعليم كمنظومة تحكم ذاتى ليس فى خاتمة المطاف وليس هو الحكم النهائى كما هو الحال فى التصور الخطى التقليدى. وإنما يكاد يكون فى جميع مراحل هذه المنظومة، ولذلك يمكن أن يصنف التقويم إلى أربعة أنواع : التقويم المبدئى Initial أو التقويم القبلى Pre - evaluation، والتقويم التكوينى Formative والتقويم التجميعى Summative والتقويم البعدى Post- evaluation

١- التقويم المبدئى أو القبلى:

التقويم المبدئى أو القبلى يتم قبل تقديم البرنامج التعليمى أو التدريبى بالفعل، وذلك لتحديد نقطة البداية الصحيحة للتعليم أو التدريب. وفيه يتم تحديد ما يتوافر فى المتعلم من خصائص (أو سمات) ترتبط بموضوع التعلم، وتحديد المشكلات الخاصة

(*) تم الإعتماد فى هذا الجزء علي ما ورد فى: فؤاد أبو حطب - القدرات العقلية ١٩٩٦ .

التى قد توجد لدى الدارس فى علاقته بالتعلم السابق والمرتبطة بموضوع التعلم الجديد، أى أن هذا النوع من التقويم يقوم بدور تشخيصى هام. ومن أمثلة ذلك ما يلى:

أ- تشخيص نقائص المتطلبات السابقة للبرنامج، وقد يتطلب الأمر تعديل البرنامج حتى يمكن للدارس اكتساب هذه المتطلبات السابقة Prerequisites ويعد هذا أحد مجالات التربية التعويضية.

ب- تشخيص مدى الاتفاق القبلى لأهداف البرنامج: فقد نجد أن بعض الدارسين يحرزون بالفعل جميع أهداف البرنامج أو عددا كبيرا منها، وفى هذه الحالة فإنهم إما أن ينتقلوا إلى مستوى أعلى فى برنامج آخر أو تتحدد لهم نقطة بداية أخرى ملائمة ويستخدم فى هذه الحالة ما يسمى الاختبار القبلى Pre-test فإذا كان البرنامج لا يتميز بمرونة كافية تسمح بهذا الانتقال الذى يعتمد على فلسفة «التفريد» فى التعليم والتدريب فلابد من أن تتوافر بدائل تعتمد على مفهوم الإثراء Enrichment.

٢- التقويم التكوينى:

أول من إستخدم مصطلح التقويم التكوينى هو سكريفن Scriven عام ١٩٦٧ فى مجال تطوير المناهج والبرامج، وفى رأيه أن من الممارسات المعتادة أنه حالما يصل البرنامج إلى صورته النهائية فإن كل شخص مرتبط به يقدم الأدلة على حاجته إلى التعديل، ولذلك يقترح ان يتم تقويم المنهج أو البرنامج أثناء بنائه أو تجربته. وذلك بجمع البيانات الملائمة والتى يمكن الاعتماد عليها فى أى تعديل ندخله عليه.

وقد إستخدم بلوم وزملاؤه (Bloom, et, al 1971) هذا المصطلح فى الأغراض التقويمية العامة للتعليم، وليس لبناء المناهج أو البرامج وتطويرها فحسب. وأصبح معناه عندهم إستخدام التقويم المنظم خلال مسار عملية التعليم أو التدريب (بالإضافة إلى بناء البرامج بالطبع) بغرض تحسين هذه العمليات. وحيث أن التقويم التكوينى يتم خلال مراحل التعليم والتدريب المختلفة فإنه يصبح أكثر وظيفية فى تحسين المنظومة وتصحيح مسارها وصولا إلى الإتقان Mastery وتتلخص أهم أغراض التقويم التكوينى فيما يلى:

أ- تقديم المعونة للمتعلم بحيث يصل إلى مستوى الإتقان.

ب- تحسين عملية التعلم والتعليم والتدريب من خلال تحليل المنوالية الكلية

للبرامج إلى وحدات أصغر يتم تقديمها بالمعدل المناسب لكل دارس وتقويمه للتأكد من الإتقان .

ج- مكافأة أو تعزيز الدارسين تعزيزا إيجابيا علي إحرازهم للأتقان أو إقترابهم منه.

د- التغذية الراجعة المعلوماتية التى تخبر الدارس بما تعلمه وبما لايزال فى حاجة إلى تعلمه كما تخبر العناصر الأخرى فى منظومة البرامج بأوجه القصور.

هـ- تشخيص صعوبات التعلم Learning Diffculties وتحديد أسبابها فى ضوء التحليل البنائى للبرنامج.

و- توصيف الطرق العلاجية البديلة فى ضوء تشخيص مواضع الصعوبة فى التعلم وأسبابها.

٣- التقويم التجميعى:

التقويم التجميعى أو النهائى موجه نحو الحكم على مدى احراز الدارس لنواتج التعلم فى البرنامج ككل أو فى جزء رئيسى فيه وذلك بهدف إتخاذ قرارات عملية مثل نقل الدارس إلى مستوى جديد أو تخرجه أو منحه شهادة أو أجازة. ويمكن أن نلخص أغراض هذا النوع من التقويم فيما يلى:

أ- الوظيفة الرئيسية للتقويم التجميعى هى الحكم ويشمل ذلك إعطاء الدرجات والتقديرات والترتيب الخ.

ب- يفيد التقويم التجميعى فى إتخاذ القرارات العملية مثل الإنتقال من مستوى إلى آخر، والتخرج، أى الاعتراف بالمهارة أو المعرفة أو القدرة، ومنح الشهادة وهذا الاعتراف أو الشهادة أو الإجازة يجب ان ينظر إليه فى الوقت الحاضر فى ضوء النظر إلى التربية على أنها عملية مستمرة، وأنه لا يتعدى حدود فترة زمنية بعدها يحتاج الدارس إلى إعادة التعلم أو إعادة التدريب (مفهوم التعلم مدى الحياة).

ج- تعتبر نتائج التقويم التجميعى نقطة بدء ملائمة لتعلم لاحق، أى أن نتائجه قد تعتبر نوعا من التقويم المبدئى لهذا التعلم الجديد، كما قد يقوم بدور التقويم التكوينى وخاصة فى مواقف التعلم المستمر.

د- يقوم التقويم التجميعى - كغيره من صور التقويم الأخرى- بدور التغذية

الراجعة ، إلا أن ذلك يتطلب أن تكون معلومات هذا النوع من التقويم أكثر تفصيلا ولا تعتمد على مجرد حكم كلى أو تقدير عام، وأن تكون متاحة للجميع وليس أسرار «امتحانية» كما هو الحال فى معظم الأحوال.

٤-التقويم البعدى:

هذا النوع من التقويم إقترحه فؤاد أبو حطب (١٩٩٦) وهو النوع الذى يتم بعد إنتهاء البرنامج وإنقضاء فترة زمنية، قد تطول أو تقصر، على إنتهائه، وتتلخص أغراض هذا النوع من التقويم فيما يلى:

١- التحقق من مدى احتفاظ الدارسين بنواتج التعلم التى سعى البرنامج إلى تحقيقها كمنظومة.

٢- التحقق من مدى قابلية نواتج التعلم التى إكتسبها الدارسون للإنتقال إلى مواقف جديدة.

٣- تتبع خريجى البرنامج للتحقق من مدى كفاءتهم سواء كان ذلك فى برنامج من مستوى أعلى أم فى مجال الحياة المهنية والعملية.

٤- تعرف أوجه النقص فى البرنامج والتى لا تربطه بأوجه الحياة العملية، وتزويد مصممى البرامج التعليمية بمعلومات تغذية راجعة تفيد فى تطوير هذه البرامج.

٥- تعرف مدى حاجة خريجى البرنامج لبرامج تدريبية جديدة سواء على نفس المستوى (برامج علاجية أو تنشيطية) أو من مستويات أعلى.

٦- تعرف مدى حاجة خريجى البرنامج لتطوير كفاياتهم المهنية من خلال توظيف فعال لمفهوم التعلم مدى الحياة.

دور الاختبارات فى البرنامج التقويمى:

على الرغم من أن دور الاختبارت فى البرنامج التقويمى يبدو لنا متواضعا، حيث يحتل خطوة واحدة من خطوات هذا البرنامج، إلا أنه فى واقع الأمر أكبر أثرا وأشد خطرا من ذلك، ونكاد نقول أن الاختبارات تكاد تكون القاسم المشترك بين معظم مكونات البرنامج التقويمى ابتداء من الأهداف وانتهاء بالأساليب التطويرية والعلاجية،

وفيما يلى نعرض لبعض الأدوار التى تقوم بها الاختبارات فى مختلف جوانب هذا البرنامج ويشمل ذلك: الأهداف واتخاذ القرارات والإجراءات التطويرية وخاصة الإرشاد النفسى.

أولا: دور الاختبارات فى صياغة الأهداف:

تساعد الاختبارات فى اختيار الأهداف التربوية وتوضيحها على نحو غير مباشر، فهى تحفز المؤسسة التربوية على صياغة أهدافها والتعبير عنها بوضوح فى عبارات سلوكية أو إجرائية، إلا أن هذا التحديد السلوكى الواضح فى ذاته ليس ضمانا لاختيار أهداف تربوية صالحة، فعملية اختيار الأهداف ليست من العمليات البسيطة التى تعتمد على مجرد التعرف أو الحكم المباشر، وإنما هى من الأمور التى تتطلب المفاضلة بين أنظمة قيمية ومعيارية، ومع ذلك فهناك مجموعة من المعلومات التى يمكن للمدرسة أو المعلم أن يحصل عليها مما يهيىء فرصًا أفضل لاختيار أكثر حكمة للأهداف، ومجموعة المعلومات هذه كما يصنفها رالف تايلر (٤٧) هى:

١ – مجموعة من المعلومات تتعلق بالأفراد أنفسهم وقدراتهم ومعلوماتهم ومهاراتهم وميولهم واتجاهاتهم وحاجاتهم، وبالطبع ترتبط هذه الخصائص بالمرحلة النمائية التى يمر بها هؤلاء الأفراد ومطالبها فى نواحى النمو الجسمى والفسيولوجي والعقلى والانفعالى والاجتماعى والمهنى، ومراعاة هذه الخصائص فى صياغتنا للأهداف يجعل عملنا التربوى معينا للفرد على تشخيص صعوباته وعلاج مشكلاته وإشباع حاجاته ومواجهة مطالبه.

٢ – مجموعة من المعلومات تتعلق بمطالب المجتمع، ويتطلب ذلك معرفة جوانب القوة والضعف التى يؤثر بها المجتمع الحديث على عمليات التربية والتدريب عموما بحيث يكتسب الفرد من العملية التربوية المعلومات والمهارات والقدرات والاتجاهات والميول التى لها أهمية فى المجتمع والتى تساعده على التكيف مع الثقافة وتحسينها.

٣ – مقترحات المتخصصين فى مختلف ميادين المعرفة والتى يرون بها كيف يمكن لميادينهم أن تساهم فى عمليات تعليم وإعداد وتدريب الأفراد.

هذه الأنواع الثلاثة من المعلومات قد توحى بكثير من الأهداف التى قد تتعارض

فيما بينها وبالتالى يصبح من الضرورى القيام بالاختيار من بينها وهذا الاختيار يعتمد أساسا على فلسفة المدرسة أو المؤسسة التربوية أو برنامج التدريب الذى نقوم به، والتى يمكن بها تحديد الأهداف ذات القيمة الكبرى فى ضوء مفاهيم «المواطن» و«المجتمع»، كما يعتمد هذا الاختيار على نتائج الدراسات والبحوث فى سيكولوجية التعلم، وفى هذه الأحوال جميعا تلعب الاختبارات دوراً مهمًا فى تزويدنا بالمعلومات حول الأفراد مما يوحى بأهداف معينة، وبعبارة أخرى يمكن القول بأن الاختبارات تلعب دورا مزدوجا فى هذه الخطوة الأولى من عملية بناء خطط التعليم والتدريب: تستثير الهيئات التعليمية لاختيار وتحديد أهدافهم بوضوح حتى يمكن استخدام الاختبارات فى تقويمها، كما أنها ذات فائدة كبيرة فى الحصول على المعلومات التى تستخدمها عندما نتخذ قرارا باختيار أهداف معينة دون غيرها.

ويتوافر للأخصائيين فى التربية فى الوقت الحاضر مصادر مهمة للأهداف قام بإعدادها المهتمون بالتقويم والقياس النفسى والتربوى، سواء فى المجال المعرفى (٢٦) أو الوجدانى (٣٦) أوالحركى، ويمكن الرجوع إلى مصدر آخر للحصول على مزيد من التفاصيل (١٤).

ثانيا – دور الاختبارات فى اتخاذ القرارات:

كل من يتعامل مع الناس يقوم باستمرار باتخاذ قرارات فيما يختص بسلوك هؤلاء الناس، وعملية اتخاذ القرارات بشأن الآخرين موجودة فى تعاملنا الاجتماعى بعضنا مع بعض، حتى فى مواقف الحوار بين شخصين، حيث يتخذ أحد المتحدثين قرارا بتغيير الموضوع لأن الشخص الآخر بدا عليه من العلامات ما يوحى إليه بأن موضوع الحديث مؤلم له أو غير مرغوب فيه على أقل تقدير، وكذلك بالنسبة للمواقف الاجتماعية الأخرى الأكثر تعقدا. كما أن عملية اتخاذ القرارات تحتل مكانا مهماً فى الشئون الإدارية وسياسة الأفراد حيث تقتضى طبيعة العمل اتخاذ قرارات فيمن يقبلون وفى توزيعهم على تخصصات مختلفة وترقيتهم بل والاستغناء عنهم، وكذلك نجد لاتخاذ القرارات أهميتها التى نحس بها جميعا فى مجال العملية التربوية، فالمدرس يتخذ قرارا خاصًا بمدى استعداد تلاميذه للانتقال إلى موضوع دراسى جديد، والمدرسة تتخذ قرارات فيما يختص بتوزيع التلاميذ، ونقلهم بل وفى معاملة من

يصدر عنهم مشكلات سلوكية ملحوظة، أو من يسعون إلى حل مشكلاتهم لدى المتخصصين فى المدرسة مثل المشرف الاجتماعى، أو الموجه التربوى، أو رائد الفصل أو ناظر المدرسة أو أى مدرس يثق فيه التلميذ صاحب المشكلة.

وكلما كان لدى الشخص الذى يتخذ القرارات معلومات كافية وصحيحة إلى حد بعيد عن الفرد أو الأفراد الذين يتخذ قرارا بشأنهم كانت الفرصة أمامه أرحب لأن يحصل على أطيب النتائج وأنسبها، فتوفر المعلومات قبل اتخاذ القرار يساعد متخذ القرار على أن يصدر أحكم قرار ممكن.

والاختبارات، على اختلاف أنواعها نفسية أو تربوية أو طبية هى وسائل لإعطاء معلومات لا نستطيع أن نحصل عليها بغيرها، وهذه المعلومات تساعدنا فى اتخاذ القرارات،

وأى قرار يتضمن تنبؤا، ذلك لأن الاختبارات كلها تدلنا على بعض الاختلافات فى أداء أفراد معينين لحظة إجراءات الاختبارات عليهم، ولكن هذه الحقيقة تكون عديمة القيمة إذا لم تستطيع أن تتنبأ أن هؤاء الأفراد يختلفون فى أداء آخر، أو فى نفس الأداء، وفى وقت آخر، وللاختبارات التى تتنبأ أهمية فى النواحى التعليمية، وفى العلاج النفسى وفى البحث العلمى.

وتستعمل الاختبارات لاتخاذ قرارات من أنواع كثيرة، منها: انتقاء الأفراد وتصنيفهم، وتقويم طرق المعاملة المختلفة (طرق تدريس مثلا)، وقبول أو رفض الفروض العلمية، وسوف نتكلم عن كل من هذه الأنواع من القرارات فيما يلى:

١ - **الانتقاء والتصنيف:** الانتقاء عبارة عن قبول أو رفض، فعندما نعطى اختبارات متنوعة لمجموعة من الشبان المتطوعين للخدمة العسكرية أو المجندين، ثم نقرر قبول بعضهم ورفض البعض الآخر حسب نتائج هذه الاختبارات المتنوعة ومستعينين بكل ما يمكننا الحصول عليه من معلومات خاصة بهؤلاء الشبان، يكون عملنا هذا انتقاء أما التصنيف فيقوم على تحديد المعاملة التى يتلقاها المتلقى، فإذا نظرنا فيما سيتم بشأن الشبان المقبولين للعمل فى الخدمة العسكرية فنجد أن منهم من سيوزع على فرع الطيران، أو الميكانيكا أو المشاة أو غيرها ويتلقى تدريبا أو معاملة مناسبة لهذا التصنيف، وكذلك نجد الشخص المصاب باضطراب انفعالى (أو مرض

نفسى) يتحدد نوع العلاج الذى يتلقاه بناء على التشخيص المبنى على الاختبارات المختلفة، وهذا التشخيص عبارة عن تصنيف نوع الاضطراب أو المرض، ويساعد التصنيف فى التركيز على نوع المعاملة العلاجية المناسبة لهذا الشخص والتى تؤدى إلى أحسن النتائج معه.

ومن التصنيف أيضًا تعيين الأفراد على «مستويات» للعمل وليس فقط لأنواع العمل، فإذا كان من نتائج اختبارات أعطيت للتلاميذ فى اللغة أن قمنا بتوزيعهم على مستويات تتراوح بين الضعيف والتوسط والحسن، ليتلقوا تدريبا لغويا مناسبا، كل حسب مستواه، كان هذا تعيينا للأفراد على مستويات للعمل.

ومن الضرورى أن نميز بين اختبارات الانتقاء واختبارات التصنيف فإذا أعطينا اختبارات كانت نتيجتها استبعاد من عندهم اضطراب انفعالى مثلا كانت هذه اختبارات انتقاء، لأن القرارات التى تبنى على المعلومات التى نحصل عليها منها قرارات قبول أو رفض فحسب، لاتدلنا على أكثر من مدى الصلاحية أو عدم الصلاحية للقبول فى فئة أو مؤسسة ما، أما الاختبارات التى تعطى معلومات تساعد على تحديد أو التعريف بطريقة معاملتهم أو تدريبهم أو علاجهم فهى اختبارات تصنيف.

ومن المهم أن نشير إلى ضرورة الوصف الفردى المفصل فى عملية التصنيف ونعنى بهذا ضرورة الحصول على معلومات مفصلة قائمة على اختبارات متعددة ومن مصادر مختلفة عن الفرد، من ناحية خصائصه الشخصية المميزة، واتجاهاته وميوله، واجتماعيته أو طبيعة علاقاته الاجتماعية، إلى جانب نتائج الاختبارات المحددة التى تهدف إلى تصنيف معين، فمن اللازم أن تكون تحت أيدينا مثل هذه المعلومات عن الطالب الذى يرغب الدخول فى كلية الهندسة أو الطب إلى جانب نتائج اختبارات الاستعداد.

(٢) **تقويم المعاملات:** إذا كانت للاختبارات أهميتها فى اتخاذ قرارات كما بينا منذ قليل، فإن لها أهميتها أيضًا فى تقويم المعاملات، أيا كان نوع هذه المعاملات تدريبا أو تدريسا أو علاجا. ولعله يكون واضحا أن ما نعنيه بالمعاملة هنا هو الطريقة التى يعامل بها فرد أو أفراد أيا كان مجالها أو نوعها، فالاختبارات التحصيلية مثلا، أو

اختبارات التحصيل المدرسى، يمكن استخدامها لمقارنة طريقتين من طرق تدريس مادة الرياضيات فى المدرسة الإعدادية للتعرف على أفضل الطرق فى تدريس هذه المادة، ويمكن الحكم على سياسات الإشراف والأفراد ومقارنة فاعليتها عن طريق قياس الاتجاهات والروح المعنوية بين الأفراد.

(٣) **التحقق من الفروض العلمية:** كثيرا ما تستعمل الاختبارات النفسية بأنواعها المختلفة فى التحقق من الفروض العلمية التى نضعها فى مجالات علم النفس المختلفة. والفرض العلمى بصفة عامة عبارة عن توقع معين أو تحديد لعلاقة وظيفية بين متغيرات، ونستخدم الاختبارات للتحقق عن مدى صحة هذا الفرض أوخطئه، فإذا أقمنا فرضا علميا يذهب إلى أن الأطفال الذين لاينشأون فى وسط عائلى طبيعى بين والديهم وأخوتهم يتميزون بدرجة ما من الاضطراب الانفعالى بعكس الأطفال الذين ينشأون فى وسط عائلى طبيعى، فإن من الممكن التحقق من صحة هذا الفرض بإعطاء مجموعتين من الأطفال، مجموعة نشأت فى وسط عائلى طبيعى ومجموعة أخرى لم تنشأ فى هذا الوسط، اختبارات شخصية تدلنا على درجة إتزانهم الانفعالى ونقارن نتائج المجموعتين بعضها بالبعض الآخر، وكذلك إذا كان لدينا فرض علمى يقول أن الأطفال يتعلمون بدرجة أحسن إذا تعاونوا فى مواقف التعلم فإننا نستطيع التحقق من هذا بإعطاء اختبارات تحصيلية متكافئة لمجموعتين من التلاميذ يتشابهون فى كافة الخصائص ولكن مجموعة منهما تتعلم بالطريقة التعاونية بينما تتعلم الأخرى بطريقة مألوفة، ومن مقارنة نتائج الاختبارات التحصيلية للمجموعتين يمكن أن نختبر مدى صحة فرضنا، وهكذا فى سائر المجالات النفسية والتربوية.

(ثالثا) دور الاختبارات فى الإرشاد النفسى:

الإرشاد النفسى Counseling هو العملية التى يمكن بها تنظيم أو عرض المعلومات الخاصة بالفرد على نحو يساعده فى الوصول إلى حلول فعالة لمشكلات التكيف التى يعانى منها والتى تكون عادة فى المستوى السلوكى العادى أو السوى، وبهذا المعنى فإن الإرشاد النفسى جانب متخصص من جوانب الخدمات الفردية، أو برامج التوجيه بوجه عام، أما إذا اتجهنا نحو مستوى الانحراف السلوكى الشديد فإننا ندخل فى باب علم النفس الطبى أو الطب النفسى والذى يتناول الأمراض النفسية

والعقلية (كالعصاب والذهان) شديدة الخطورة سواء كانت ترجع لأسباب عضوية أو غير عضوية، وإذا اتجهنا نحو المستوى السلوكى الأكثر سواء ندخل ميدان الخدمات الروتينية الخاصة بالتعرف المهنى وإعطاء المعلومات عن المهن أو الفرص التعليمية. ولكن المشكلات الأبعد مدى من مجرد إعطاء المعلومات والأقل من المرض النفسى أو العقلى أو سوء التكيف الاجتماعى (انحراف الأحداث أو الاجرام) قد تكون مشكلات ملحة ومتنوعة وغالبا ما تكون على درجة من الصعوبة والتعقيد، والأغلب تكون ذات أثر طويل المدى، ومن هذه المشكلات ما يتعلق بالتخطيط للمهنة والتخطيط لنوع التعليم الملائم للفرد، ومشكلات التكيف الاجتماعى والإنفعالى التى يواجهها المراهقون سواء فى بداية مرحلة المراهقة أو نهايتها، والمشكلات الخاصة بصعوبات التعلم والتى تتطلب تعليما علاجيا من نوع معين.

فالمرشد النفسى الذى يتناول هذه المشكلات وأمثالها بالبحث والدراسة والتشخيص والعلاج هو واحد من مجموعة كبيرة من المتخصصين فى الخدمات الفردية منهم الأطباء والأطباء النفسيون والأخصائيين الاجتماعيون والمتخصصون فى عيادات الأطفال، والإرشاد النفسى مثله مثل تلك التخصصات لاينشط فى فراغ ولكن تحدده البيئة وظروف المجتمع التى تؤثر دون شك فى وسائل المرشد النفسى وأنشطته وأهدافه، فالنظام التعليمى عموما يتضمن أنماطا من المناهج الخاصة التى تحددها ضرورات اجتماعية أحيانا أخرى، وتتمثل الضرورة القانونية فى أن الطبيب لايمكن أن يمارس هذه المهنة إلا بعد التخرج فى كلية الطب، على حين أن إصرار أصحاب الأعمال على التخرج فى أقسام الصحافة وإدارة الأعمال مثلا يمثل الضرورة الاجتماعية عندنا، كما أن الأسر المصرية حين ترسل أبناءها إلى المدرسة إنما تعتبر عملية التعليم فى ذاتها صورة متخصصة للتأمين ضد البطالة، وبالتالى فإن الآباء يلحون على أبنائهم أن يبذلوا قصارى جهودهم لمواصلة تعليمهم.

أما النظام التعليمى المصرى نفسه فإنه يؤكد أساسا الدرجات المدرسية ومجموعها ونظام القبول حسب مكاتب التنسيق التى أصبحت من الثوابت الرئيسية فى نظامنا التربوى، والتى امتدت من القبول بالمدارس الإعدادية حتى القبول بالكليات والمعاهد العليا.

وإذا كان لوظيفة الإرشاد النفسى أن تحقق أغراضها فى مجتمعنا فلابد أن تعمل خلال المحددات التى تضعها هذه الاتجاهات والمطالب والتقاليد الاجتماعية، وسواء اتفقنا أو اختلفنا معها فإنها تمثل الإطار المرجعى القائم الذى يعمل المرشد النفسى فى حدوده، تماما كما يكون على الأطباء أن يتناولوا مشكلات الصحة الصناعية لصناعة معينة فى منطقة معينة، هذا من حيث الشكل الذى يحدد خدمات الإرشاد النفسى. أما من حيث المضمون فإن هذه الخدمات وما يتعلق بها من تقويم وقياس تتأثر بأهداف المؤسسة (تربوية كانت أم مهنية) وأغراضها، فإذا كان الهدف الرئيسى هو الانتقاء أو النصح أو التزكية أو الاستشارة، فإن معنى ذلك أن إجراءات الإرشاد النفسى وعمليات التقويم والقياس المرتبطة بها تختلف عما لو كان الهدف الأكبر أكثر عمومية وأكثر اهتماما بنمو الأفراد وتحسين مستوياتهم.

والمرشد النفسى هو فى العادة أخصائى كلينيكى يتعامل مع فرد واحد فى المرة الواحدة، ولذلك فهو عادة ما ينزع إلى أن يكون أكثر اهتماما بجوانب السلوك الفردى، ملتزما بمعايير متعددة الأبعاد لتكيف الفرد، وهو لايهتم بالنجاح كما يقاس بالتحصيل المباشر فحسب ولكن بالنجاح كما يقاس بالرضا الفردى والبقاء والاستمرار فى المدرسة أو المهنة، كما أنه لايحاول أن يتنبأ بنتائج برنامج معين فى القراءة العلاجية مثلا فحسب ولكن بالتكيف الاجتماعى للفرد كما يتأثر بالخبرات الاجتماعية والنمو، ومن واجبه ألا يركز على الصراع العاطفى المباشر فحسب، ولكن على آثار هذا الصراع فى النمو العاطفى والانفعالى بوجه عام.

ولكن ما هو الدور الذى يقوم به التقويم والقياس فى خدمات الإرشاد النفسى؟

يمكن القول بأن علاقة التقويم بالإرشاد النفسى هى علاقة الوسيلة بالغاية، أى المساعدة على التعرف على نواحى القوة والضعف سواء كانت داخل الفرد أو بين الأفراد. بعبارة أخرى يفيد التقويم والقياس فى جمع المعلومات عن الأفراد الذين يعانون من المشكلات الفردية الخاصة بالتكيف والتوجيه والنمو، كذلك مشكلات الجماعات من حيث التحصيل والانتقاء أو التوجيه المهنى والتربوى والتصنيف والتقسيم سواء فى الميادين التربوية أو المهنية على مجرد جمع المعلومات وإنما يمتد إلى زيادة استبصار المرشد النفسى وفهمه لسلوك الأفراد ولسلوكه هو أيضًا حتى

يستطيع الوصول إلى التشخيص المناسب، كما يسمح التقويم والقياس، وخاصة إذا توفرت فيه شروط معينة سنناقشها فى هذا الكتاب، بالوصول إلى تنبؤ أدق مما لو استخدمنا أية وسائل لا تتوفر فيها تلك الشروط.

والواقع أن التشخيص والتنبؤ المعتمدين على الوقائع لها أهمية خاصة فى حياتنا الاجتماعية عموما، فمثلا لايمكن الاعتماد على أى برامج للانتقاء أو التوجيه إلا إذا سعت هذه البرامج بالتنبؤ الكمى الجيد، وعلم الطب يعتمد أساسا على التشخيص من ناحية وعلى القيمة التنبؤية للمقاييس الفسيولوجية من ناحية أخرى، كما أن كلا من الطب النفسى والخدمة الاجتماعية قد أفاد كثيرا من المعلومات الكمية التى نجمعها عن سلوك الأفراد.

وهذا المنطق يصدق بالطبع على خدمات الإرشاد النفسى. وهناك اتجاه يرى استخدام الاختبارات بأنواعها مسألة أساسية فى العلاج والإرشاد النفسى، ذلك لأنها أدوات يتحدد فى ضوئها تشخيص نوع الصعوبة الذى يعانى منها العميل، ومنها يستطيع المعالج أو المرشد النفسى أن يكون فكرة أقرب إلى الدقة عن هذا العميل، كما أن الاختبارات - وخاصة اختبارات الشخصية - تساعد فى التعرف على مدى التقدم الذى يحرزه الشخص فى أثناء عمليه الإرشاد النفسى أو العلاج النفسى.

ويمكن أن نلخص الأسباب التى من أجلها يؤكد أصحاب هذا الاتجاه أهمية الاختبارات فى عملية الإرشاد النفسى فيما ذكره بوردين (٢٧: ٢٦٢ - ٢٧٩) وتايلور (٤٧) كما يلى:

١ - تمد الاختبارات المرشد النفسى بمعلومات معينة يحتاج إليها فى عمله مع الشخص وتساعده فى مهمته معه.

٢ - تمد الشخص أو العميل بمعلومات معينة يحتاج إليها وخاصة فى اتخاذ قرارات معينة.

٣ - قد تستثير الاختبارات العميل لاستكشاف ذاته.

٤ - أن العميل قد يكتسب استبصارا وفهما لنفسه وهو يقوم بأداء الاختبار.

إلا أن هناك اتجاها آخر ينظر إلى استعمال الاختبارات فى العلاج والإرشاد

النفسى نظرة مغايرة، إذ يتردد أصحاب هذا الاتجاه فى الاعتماد كلية على منهج التقويم والقياس، وهذا التردد يعتمد على حجتين أساسيتين: أولاهما أن تأكيد المنطق الكمى والطبيعة العقلانية لعملية الوصول إلى استبصار بموقف العميل قد يؤدى إلى تضليلنا، لأن ما نعتبره معلومات موضوعية محايدة هو فى حقيقة الأمر مرتبط بالكيان الانفعالى والعاطفى للعميل، وثانيهما أن معرفتنا الراهنة بالشخصية وأدوات تقديرها جميعا لازالت فى مراحلها الأولية المبكرة.

وهذا الاتجاه يمثله فى الوقت الحاضر كارل روجز Carl Rogers (٤٢، ٤٣) وتلاميذه حيث يركزون الاهتمام على الجوانب العلاجية فى الإرشاد النفسى ويتبعون طريقة خاصة من طرق العلاج تتميز بعدم التأكيد على الاختبارات لأنهم لا يقيمون وزنا كبيراً لعملية التشخيص والتصنيف فى العلاج، وجوهر طريقتهم هو أن الفرد مركز التقويم، أى أن التقويم يأتى من داخل الفرد، وأن نموه وشفاءه أو حل المشكلة التى يعانى منها إنما يأتى من قوى داخلية تنبعث منه عندما يوفر له المعالج الجو المطمئن الخالى من أى نوع من أنواع التهديد للذات أو التقويم للفرد، هذا الجو الذى يتحقق عندما يدعم المعالج عند العميل باستمرار وبعمق أنه يتقبله ويثق به ويهتم به باعتباره إنسانا فى أزمة وجود أو أزمة حياة، ويساعده على التغلب عليها ليستمر نموه.

هذا بإيجاز شديد هو الجوهر الذى يدور حوله العلاج المتمركز حول العميل، وفى ضوئه نفهم الأسباب التى جعلت الاختبارات عند أصحاب هذا الاتجاه لا تحتل مكانا مهماً فى العلاج أو الإرشاد النفسى، ويمكن أن نلخص هذه الأسباب فيما يلى:

(١) المجال الذى ينشط فيه الفرد هو على درجة كبيرة من الاتساع وأن مواقف الأداء الفعلى عديدة ومتغيرة بحيث لايمكن تمثيلها فى أى موقف اختبارى (الذى هو فى ذاته بحكم تعريفه ليس إلا عينة من السلوك).

(٢) أن تأثير المتغيرات والعوامل التى يصعب أو يستحيل قياسها فى الوقت الحاضر قد تكون على درجة كبيرة من القوة والأهمية، مثل عوامل الحوافز والدوافع، إلى حد أن الصفات السلوكية التى يمكن قياسها - إذ قورنت بما لايمكن قياسه - قد تكون أقل أهمية بالنسبة للنشاط الواقعى فى الحياة (٢٢).

(٣) الاختبارات ليست إلا عملية مقارنة بين العميل وآخرين فى جانب من الجوانب، بينما المشكلة التى يعانى منها العميل هى فى حقيقتها مشكلة ذاتية نابعة من إدراكه الخاص بالمواقف والأشخاص، ومن إدراكه لنفسه، ومن المعانى التى يخلعها على هذه المدركات جميعا، ولايمكن أن نفهم العميل ومشكلته إلا فى ضوء هذه الذاتية.

(٤) توحى الاختبارات إلى العميل أن دور المعالج أو المرشد النفسى هو أن يقوم بعمل شىء له (أى للعميل) يحل مشكلاته وينهى معاناته، ومن ثم يتخذ هذا العميل دوراً سلبياً ينتظر فيه ما سيعمل له بدلا من أن يحس مسئوليته فى علاج نفسه متعاونا مع المعالج أو المرشد النفسى، الذى لا يزيد دوره عن كونه - كما عبر روجرز - الأنا الأخرى للعميل، هذا الاتجاه السالب الذى تخلقه الاختبارات لدى العميل يجعله يتوقع «وصفه» لمشكلات، ويصبح من أشق الأمور على المعالج أو المرشد النفسى نقل مسئولية الموقف واتخاذ القرارات إلى العميل نفسه.

(٥) تحمل الاختبارات أيضاً إلى العميل شيئا من البعد بينه وبين المعالج أو تسبغ على العلاقة بينها صفة الرسمية أوتوحى بالعلو من جانب المعالج على العميل. وهذا النوع من العلاقة إذا جاز وجوده فى مجال الطب والأمراض الجسمية فإنه غير مناسب فى علاقة المساعدة فى مجال الإرشاد والعلاج النفسى.

(٦) قد يستعمل العميل الاختبارات عند أخذها وفى تفسيرها والتحدث حولها ستارا أو دفاعا أو هروبا من المشكلات الحقيقية التى ينبغى أن يواجههنا مما يعطل سير عملية العلاج أو الإرشاد النفسى.

بعض الوظائف التربوية للتقويم النفسى:

يؤدى التقويم النفسى وظائف عديدة فى مختلف مجالات الحياة كالصناعة والإدارة والتربية والعلاج النفسى وغيرها، على نحو يدل على أهميته وقيمته للباحث النفسى فى هذه المجالات ونعرض فيما يلى بعض هذه الوظائف فى مجال بذاته وهومجال التربية، ويمكن للمهتمين بالمجالات الأخرى أن يجدوا بسهولة ويسر وظائف التقويم النفسى واضحة لهم.

(أولا) دور التقويم فى زيادة الدافعية للتعلم:

يقوم التقويم بالوظائف الثلاث الرئيسية للدوافع فى التعلم وهى:

(أ) وظيفة التنشيط energizing أى زيادة المستوى العام للنشاط والجهد المبذول، ويمكن التدليل على ما تحدثه الامتحانات السنوية والعامة من ازدياد النشاط والجهد فى فترات العمل الشديد التى يقضيها طلاب المدارس الثانوية والجامعات قبيل الامتحانات، وكثيرا ما تكون هذه الامتحانات فى معظم المدارس خاتمة المطاف. ويكون للنجاح والفشل فيها أهمية كبيرة، كما تحدد هذه الامتحانات عموما متى يتعلم التلاميذ وماذا يتعلمون وكيف يتعلمونه، ولذلك فإن هذه الامتحانات إذا لم تقم بقياس أهداف المنهج أو المقرر قياسا حقيقيا وصادقا فإننا نشك كثيرا فى قيمة الدافعية التى تستثيرها.

ويمكن القول من الوجهة المنطقية البحتة أنه ما دام التقويم يزيد الجهد المبذول فإن تكرار إجراءات التقويم يؤدى بالتالى إلى زيادة الجهد الكلى للتلاميذ ، ولكن من المؤكد أن التكرار المناسب للتقويم يعتمد على طبيعة الأدوات التى تستخدمها، وطرق التدريس، وما إذا كان التلاميذ يصححون اختباراتهم أو اختبارات غيرهم، والمادة الدراسية، وقدرة التلاميذ بوجه عام. ولاشك أن إجراء التقويم بكثرة يقلل نسبيا من أهميته، فالامتحانات اليومية مثلا، قد لايكون لها أى أثر منشط يزيد عن أثر الواجبات المدرسية اليومية، ويلخص كوك (٢٨) نتائج البحوث التى أجريت على أثر التقويم فى الدافعية فى أنه عندما تعطى كل أسبوع الاختبارات التحصيلية فى المواد الدراسية فى مستوى الجامعة ثم تناقش النتائج وتستخرج الأخطاء الفردية، وعندما تتم صياغة أسئلة الامتحان النهائى على صورة تشبه تلك الاختبارات فإن تحصيل الطلاب من ذوى المستوى المنخفض من القدرة يكون أعلى منه فى حالة التقويم الأقل تكرارا، أما الطلاب من ذوى المستوى المرتفع من القدرة فقد يتخلفون ما لم يكن فى كل اختبار أسئلة على درجة مناسبة من الصعوبة تتحدى قدرتهم، كما أنهم يفيدون أكثر من المادة الإضافية التى يمكن للمعلم أن يدرسها إذا كرس وقتا أقصر للتقويم والعمل العلاجى بوجه عام. ولكن هل يمكن أن تعزى آثار التقويم المتكرر فى تحسين تحصيل التلاميذ الضعاف إلى وظيفة التنشيط هذه؟ الواقع أن معظم الباحثين يرون أن من

المحتمل أن هذه الآثار ترجع إلى توجيه التعلم واختيار الاستجابات الصحيحة (وهما الوظيفتان اللتان سنتحدث عنهما فيما بعد) أكثر من زيادة الطاقة والنشاط والجهد التى تستثيرها الامتحانات.

وقد تناول الباحثون بالدراسة أيضاً أثر معرفة نتائج الامتحانات (والتقويم عموما) فى زيادة الدافعية للتعلم وأكدت نتائجهم أن معرفة النتائج أفادت التعلم بمقدار النجاح الذى تتضمنه النتائج، وقد أفاد التلاميذ من ذوى المستويات المرتفعة أكثر من غيرهم أما أولئك الذين أحرزوا تقدما بطيئا فلم تستثرهم معرفة النتائج إلا استثارة ضئيلة للغاية.

وقد درست هرلوك (فى :١٤) أثر المدح العام أو اللوم العام فى التعلم، وأكدت نتائجها أن المدح أكثر استثارة للتعلم من الذم، وهو أشد تأثيرا فى التلميذ الأقل نضجا والأقل قدرة، أما الذم فكان أكثر فائدة بصفة عامة مع التلميذ الناضج والقادر وقد لوحظ أن كلا من المدح والذم أكثر استثارة للتعلم من موقف الحياد أو التجاهل التام.

وخلاصة القول: أن الأثر المنشط للتقويم يعتمد أساسا على درجة نجاح التلميذ، فالذى يحصل على درجات عالية وبالتالى ينجح ويتقدم ويلقى المديح تكون استثارته أيسر من التلميذ الذى يحصل على درجات منخفضة وما تستتبع ذلك من فشل ورسوب وذم ولوم، وهذه النتيجة ذاتها تؤكد ضرورة وضع أهداف تربوية معقولة تتضمن على الأقل نجاحا جزئيا وممكنا للجميع.

(ب) وظيفة التوجيه directive ، وبها يتم توجيه سلوك المتعلم ونشاطه إلى المسالك المرغوب فيها. فإذا كانت نتائج التقويم يمكن بها الحكم على المدارس والمعلمين والتلاميذ فإن طبيعة أدوات التقويم تحدد إلى حد كبير نوع العملية التربوية، فإذا كانت الاختبارات تتضمن أحكاما كيفية مهمة فإنها تفيد كثيراً فى تحديد الأهداف والطرق التربوية، وإذا كانت الاختبارات تعتمد على مضمون تقليدى بحيث تدور معظم أسئلتها حول الحقائق بغرض تحديد مقدار ما حفظه التلاميذ من مادة الكتاب المدرسى، فإن أثار هذا النوع من الاختبارات تختلف عن الاختبارات التى تقيس مهارات الدراسة وقدرات حل المشكلة، لأن الأخيرة تؤكد المفاهيم وتطبيقاتها فى المواقف الجديدة وتوضح الأهداف المهمة وتركز الانتباه على الهدف الأقصى لعملية

التربية وهو التعلم الدائم، ويمكن تلخيص الفرق بين نوعى الاختبارات فى أن النوع الأول يجمد المنهج المدرسى والنوع الثانى يذيبه.

ولاشك فى أن التقويم يؤثر فى إجراءات الدراسة وأهدافها كما سبق أو أوضحنا. ففى إعداد امتحانات من نوع المقال يكون على التلاميذ تحديد المادة وتنظيمها فى وحدات كبيرة وتأكيد العلاقات والاتجاهات والاستجابات الشخصية الخاصة بالتلميذ الفرد، أما فى إعداد الاختبارات الموضوعية التى تعتمد على الحقائق فإنها تتناول عينة من السلوك أشمل من اختبارات المقال ولكنها تتطلب من التلاميذ تأكيد التفاصيل والأسماء والتواريخ ونتائج التجارب ، وقد لوحظ أيضًا بعض الاختلاف فى طريقة إعداد التلاميذ لأنفسهم عند الإجابة على الأنواع المختلفة من الاختبارات الموضوعية، فمثلا عند الاستعداد لاختبارات التكملة فإن الطالب يحاول إتقان أسلوب الكلمة فى مقابل الكلمة، أما فى حالة الاستعداد لاختبارات الصواب والخطأ أو الاختيار من متعدد بوجه عام فإن التلميذ يهتم بالتعريفات والحقائق التفصيلية، ولاشك أن هذا من العوامل التى تحدث فروقا فيما نتعلمه وما نحفظه لفترة من الزمن وفى القدرة على أداء أنماط مختلفة من الإختبارات، وكلها ترتبط تجريبيا بنوع الاختبار الذى يستعد الطالب للإجابة عنه.

(جـ) وظيفة الانتقاء selective، وتعنى هذه الوظيفة تحديد الاستجابات التى سوف يتم تثبيتها وبقاؤها عند التلاميذ والاستجابات التى سوف تحذف، وترتبط هذه الوظيفة بالجانب التشخيصى للتقويم، إلا أن الفرق بين الانتقاء والتشخيص هو أن الاختبار التشخيصى يركز الاهتمام على تشخيص الصعوبات كما يقترحها المعلم أو الأخصائى النفسى المدرسى، أما ما نتحدث عنه الآن فهو أن التلاميذ والمعلم والأخصائى النفسى جميعا يمكنهم الحصول على مقدار كبير من المعرفة من أخطاء التلاميذ (وأسبابها إذا كان هذا ممكنا) فى موقف الاختبار، وينتج أكبر مقدار من التعلم من موقف الاختبار ذاته حينما يسمح للتلاميذ بتصحيح أوراق إجاباتهم ثم يلى ذلك مناقشة الأخطاء ثم العمل العلاجى، وتؤكد البحوث التجريبية أنه كلما كانت معرفة التلاميذ لجوانب الصواب والخطأ فى استجاباتهم وأسبابها معرفة مباشرة فى موقف الاختبار فإن ذلك يؤدى إلى تثبيت الاستجابات الصحيحة أوالسلوك المرغوب فيه وحذف الأخطاء واستبعادها.

(ثانيا) : دور التقويم فى تصميم المواقف التعليمية :

يعد مبدأ الفروق الفردية من المبادئ الأساسية فى علم النفس الحديث، وهو يعنى من الوجهة التربوية التباين فى جماعات الأعمار الزمنية والفرق الدراسية فى الذكاء والقدرات والتحصيل والاتجاهات والميول وسمات الشخصية المختلفة، وأى رأى يعتبر محض الفرقة الدراسية مرحلة لها خصائصها المحددة التى تنطبق على جميع الأفراد المنتسبين لها رأى خاطئ ويعطل جهود المؤسسة التربوية وسعيها لإشباع حاجات التلاميذ كأفراد. وأى افتراض حول وظيفة المعلم على أنه متخصص فى فرقة دراسية معينة إنما يلغى إمكانية نجاحه فى التعامل مع التلاميذ الذين قد يشذون عن المستوى الافتراضى لهذه الفرقة، وأى جهد من جانب السلطات التعليمية لجعل المقرر الدراسى جامدا ثابتا وعاما على جميع تلاميذ الفرقة الدراسية لايضع فى الاعتبار أن هؤلاء التلاميذ ليسوا جميعا على الدرجة نفسها من القدرة على أداء ما يطلب منهم بنجاح.

وقد أجريت عدة محاولات لحل مشكلة الفروق الفردية فى مواقف التعلم نذكر ما يلى:

١ - **مراعاة القدرة العقلية العامة:** من الافتراضات التى وجهت الممارسات التربوية فى الماضى أن الجماعات المدرسية تكون أكثر تجانسا فى التحصيل إذا تجانست فى القدرة العامة ولذلك كانت تستخدم طرق تربوية مقننة مثل توحيد المقررات والامتحانات بحيث تتفق مع هذا المستوى العام ويتضمن هذا الافتراض مسألتين أولاهما أن التباين داخل الفرد لايعتد به، وثانيهما أن جميع الخصائص السلوكية التى تهتم بها المدرسة ترتبط فيما بينها ارتباطا عاليا وتنتظم حول عامل يحدد المستوى العام لكفاءة الفرد. والواقع أن الأدلة العلمية الحديثة تدحض هذا الافتراض وتضمياته، وقد جاءت هذه الأدلة من مصادر عدة أشهرها البحوث الحديثة فى القدرات العقلية التى أكدت أن الاعتماد على القدرة العقلية العامة وحدها لاينتج مجموعات متجانسة تجانسا كاملا(١٢).

٢ - **مراعاة مستويات التحصيل السابقة:** من ناحية أخرى فإن المعلمين ورجال الإدارة التعليمية والمسئولين التربويين يرجعون التفاوت الكبير فى مدى القدرة لدى المتعلمين إلى عدم الدقة فى إجراءات نقلهم من المستويات التعليمية وهم فى هذا الصدد يطالبون بأن تكون للفرقة الدراسية الواحدة مستويات محددة يجب أن يصل

إليها التلاميذ قبل انتقالهم إلى فرقة أعلى، وعلى الرغم من جودة هذه الفكرة إلا أنها صبت أحيانا فى القالب التقليدى حيث يحدد المقرر المدرسى ما يجب أن يتعلمه الجميع فى كل مادة فى كل فرقة دراسية وأى الكتب المدرسية تستخدم، وأى صفحات أو فصول يجب أن تعطى فى وقت محدد من السنة، وتكون محصلة الأمر أن يتعلم التلميذ المعارف المتضمنة فى هذه الكتب مثل: تهجى الكلمات، المسائل النموذجية فى الحساب، الأسماء التواريخ، أسباب الحروب ونتائجها، الدول وعواصمها، والمدن وخصائصها ... إلخ.

والسؤال الآن هو: على توجد بالفعل مستويات للتحصيل لايتباين فيها التلاميذ بحيث تؤدى إلى تحقيق التجانس داخل الجماعات المدرسية، للإجابة على هذا السؤال يمكن صياغة القاعدة العامة الآتية التى نستخلصها من نتائج البحوث العديدة التى أجريت فى هذا الميدان.

«إذا كانت الاستجابات التى نتعلمها بسيطة والأهداف التى ننشدها محددة بحيث يمكن لأكبر عدد من التلاميذ أن يصلوا إليها خلال فترة التعلم فإن تباين المجموعة يتناقص، أما إذا كان العمل على درجة من التعقد والتركيب والأهداف غير محددة بحيث لايمكن أن يصل إليها إلا التلاميذ الأقدر فإن تباين المجموعة يتزايد».

ومعنى ذلك أن الرأى القائل بأن التلاميذ يمكنهم أن يصبحوا أكثر تجانسا إذا استخدمنا أساليب التدريس التى تؤدى بهم إلى مستوى واحد مقنن إنما يفترض محدودية الأهداف وضيق نطاق التدريس فحتى يمكن أن تكون المهام فى متناول التلاميذ الأقل قدرة يلجأ المعلم إلى أساليب التعلم الروتينى، كما أن الكتب المدرسية تكون فى العادة من مستوى أدنى، والمواد التى يتم اختيارها للمنهج تكون من النوع الذى يمكن أن يتقنه حوالى ٨٠٪ من التلاميذ، ويتم تدريب التلاميذ على الأسئلة وإجاباتها، مثل أسئلة الامتحانات، وعادة ما يكون معظم وقت المعلم مكرسا للتلميذ بطىء التعلم وفى هذه الحالة عندما يقاس تقدم التلميذ فى فترة معينة فى ضوء الأهداف المحدودة فإن التلاميذ ذوى الدرجات المنخفضة فى البداية يحرزون أعلى تقدم ممكن ومعنى ذلك أن هناك علاقة سالبة بين الدرجات المبدئية وبين درجات الاستفادة أو الكسب فى نهاية التعلم، كما أن العلاقة بين الاستفادة فى هذه الحالة

والذكاء علاقة منخفضة وكثيرا ما تكون سالبة، وهذه النتائج تؤكد أن مثل هذه الممارسة لا تستثير التلاميذ من ذوى القدرة العالية أو المتوسطة إلا قليلا إن لم تهملهم إهمالا تاما.

وقد تكون أخطر النتائج المترتبة على ذلك أن ما يتعلمه المتعلمون قليل القيمة، ضعيف الأثر، ضيقا فى مدى الاحتفاظ به وتستخدم فيها أسئلة الحقائق والتفاصيل والمعلومات المحدودة التى تدور حول أذكر، وصف، وعرف، ومن، وماذا، ومتى، وأين، ولايستخدم التلاميذ فى الإجابة عليها إلا الذاكرة، وتكون آثار التعلم فى هذه الحالة مؤقتة وخاصة مع استخدام الذاكرة الصماء، لأن ما ينساه التلميذ فى هذه الحالة تكون ما بين ٤٠٪ خلال ثلاثة أشهر و٨٠٪ خلال ثلاث سنوات ويسمى باكنجهام (فى ١٢) ذلك «أعظم فقد تربوى»، ومنحنيات النسيان لمعظم المواد الدراسية فى هذه الحالة تقترب كثيرا من منحنيات نسيان المواد عديمة المعنى.

أما إذا كان التعلم يتضمن عمليات حل المشكلة وغيرها من العمليات العقلية العليا فإن أثاره تكون أكثر دواما، أى أنه ما لم يعمل التعلم على تمايز وتكامل الاستجابات القديمة والجديدة فى نمط سلوكى منظم لاتكون لهذا التعلم قيمة أو استمرار، وهذه النتائج تهمنا كثيرا فى الوقت الحاضر مع انتشار مفهوم «التعلم للاتقان» وحتى لايتحول هذا الشعار النبيل إلى محض وصول إلى مستوى متواضع من القدرة.

خلاصة القول أن حل مشكلة الفروق الفردية فى مواقف التعلم لايكون بتحقيق تجانس التلاميذ فى القدرة العامة أو باستخدام الطرق المقننة وإنما بالإجابة على سؤال أهم هو كيف تم التعلم وكيف صممت مواقفه؟ فإذا سادت التعلم العمليات العليا وكانت أهدافه من العمومية والاتساع والارتفاع أدى ذلك إلى تحقيق نتائج أكثر استمرارا ودواما وإلى استثارة المتعلمين لبذل قصارى جهدهم فى التعلم فإننا بذلك نصل بالمتعملين إلى الحد الذى يظهر قدراته وعلى نحو فيه التباين الإبداعى أكثر مما فيه من التجانس الجامد.

(ثالثا) دور التقويم فى تصنيف المتعلمين:

لجأ بعض الباحثين فى حلهم لمشكلة الفروق الفردية إلى افتراض أن التباين داخل الفرد والتباين بين الأفراد يدعوان إلى ضرورة مقابلة الحاجات الفردية، ويعنى

ذلك أن نقبل حقيقة وجود التباين الواسع فى السلوك الإنسانى.

ولمواجهة هذا المطلب فى المدرسة الابتدائية مثلا اهتم الباحثون بخصائص نمو التلاميذ من ناحية، وتكيفهم المناسب للثقافة التى يتأثرون بها من ناحية أخرى، ففى عملية التفاعل الاجتماعى العام خلال سنوات التكوين تكون السمات الأكثر أهمية فى تحديد عضوية الطفل ومكانته فى الجماعة هى العمر الزمنى والنمو الجسمى العام، وإذا أضفنا إلى ذلك النمو الاجتماعى العام، تكون هذه السمات الثلاث فى رأى المهتمين أساس التصنيف والتجميع التربوى، إلا أن ما يجب ملاحظته أن العمر الزمنى لا يرتبط ارتباطا كاملا بالنمو الجسمى والاجتماعى للطفل، ولهذا تعطى الأولوية لمستوى النمو لا للعمر الزمنى وبخاصة النمو الجسمى والاجتماعى، حتى يمكن للطفل أن يعيش ويعمل فى مجموعة ينتمى إليها تقبله ويقبلها.

ويضاف إلى عامل النمو (وخاصة النمو الجسمى والاجتماعى) فى التصنيف باعتباره العامل الأولى والأساسى عامل آخر هو مستوى التحصيل النوعى فى المجالات المختلفة تبعا لقدرات التلاميذ فى مختلف المواد الدراسية. ويجب أن نتذكر دائما أنه حتى لو تم التجميع والتصنيف فى المدرسة الابتدائية بهذه الطريقة فلابد أن يتركز اهتمامنا على التلميذ الفرد. وبالطبع فإن ملاحظات المعلم لتلاميذه وحكمه عليهم يكون لها دائما أهمية قصوى فى موقف التعلم.

دور الحاسب الألكترونى فى التقويم النفسى والقياس:

يتزايد فى الوقت الحاضر استخدام الحاسب الالكترونى (الكمبيوتر) فى المجالات المختلفة لعلم النفس سواء داخل المعمل أو خارجه، ولعل أكثر هذه المجالات استخداما للحاسب الألكترونى مجال التقويم النفسى. ومن المتوقع أن يزداد هذا الاستخدام شيوعا وانتشارا فى المستقبل.

وتمتد الاستخدامات الراهنة للحاسب الألكترونى على نطاق عملية التقويم النفسى كلها بدءا من بناء الإختبار وتطبيقه وتصحيحه، إلا أن أكثر هذه الاستخدامات وضوحا هى عملية تصحيح الاختبارات وتحليل وتفسير نتائجها. وقد ظهر هذا الإستخدام فى بداية دخول الحاسبات فى مجال القياس وبصفة خاصة فى تطبيق الإختبارات الجمعية مثل اختبارات الذكاء، الإستعدادات والشخصية.

وتتوافر لدى ناشرى الاختبارات فى الوقت الحاضر إمكانات تزويد الباحثين بنتائج الاختبارات على أى صورة يشاءون ابتداء من نتائج المؤسسات وحتى نتائج الأفراد، وذلك بعد تصحيح أوراق الإجابة التى عادة ما تتخذ صورة معينة تصلح للمعالجة بواسطة الكمبيوتر، ويستخدم الحاسب الالكترونى أيضاً بسهولة ويسر فى عمليات تحليل الأسئلة فى الاختبارات وانتقاء الملائم منها كما يمكن المقارنة بين المعايير المحلية والمعايير القومية للاختبار، ويستطيع الحاسب الالكترونى أيضاً أن يزود المدارس والمؤسسات المختلفة بنتائج الاختبار لتقوم بدورها فى التغذية الراجعة. وقد يسر ذلك على المعلمين الأخصائيين أن يستثمروا هذه النتائج فى تحسين برامج التعليم والتدريب.

ويتزايد فى السنوات الأخيرة استخدام الحاسب الالكترونى فى إعداد اختبارات التحصيل واستخدامها داخل المؤسسات التعليمية ومع ظهور الحاسبات الالكترونية المصغرة تتوافر فى الوقت الحاضر أعداد منها فى معظم المدارس الاوربية والأمريكية، تفيد المعلمين فى سرعة تصحيح الاختبارات التى يقومون بإعدادها، ناهيك عن الحاسبات الألكترونية الكبيرة التى تستخدم فى تصحيح وتحليل نتائج الاختبارات ابتداء من المستويات المحلية وحتى المستوى القومى العام.

كما ظهرت بعض البرامج التفاعلية بين الإختبار والمفحوص ومساعدته فى اتخاذ القرار والتخطيط له وذلك من خلال تخزين بعض المعلومات عن المفحوص وأداءاته فى اختبارات مختلفة والإجابة عن بعض الاسئلة، ومن تفاعل هذه المعلومات يمكن المساعدة فى اتخاذ قرار معين مثل اشتراك المفحوص فى دورة تدريبية، أو اختيار عمل معين.. الخ ومن هذه البرامج System of Interaction Guid (SIGI) ance information وقد تم تطوير هذا البرنامج واستخدم فى بعض الجامعات والكليات وسوق العمل أو تغيير طبيعة العمل أو الترقية (انستازى وأربينا : ٢٤ أ).

وبالرغم من التطور الهائل فى استخدام الحاسبات فى ميدان القياس والتقويم التربوى فقد ظهرت بعض السلبيات لهذا الإستخدام مما تطلب وجود ارشادات محددة لمستخدمى هذه البرامج وبصفة خاصة عند تفسير الدرجة التى يحصل عليها المفحوص.

وللتحقق من الدرجة التى يحصل عليها المفحوص عند تطبيق الاختبار وتصحيحه بالطريقة التقليدية، والدرجة التى يحصل عليها ذات المفحوص عند استخدام الحاسب لابد أن يكون الإختباران موزونان (Equated) لأن المعالم السيكومترية للإختبارين قد تختلفا فى الصدق والثبات وأيضا المعايير، كما يجب أيضا مقارنة الأدائين التقليدى والمحوسب لمجموعات كبيرة ومختلفة من المفحوصين قبل الإعتماد على النتائج المستمدة من الاداء المحوسب وبصفة خاصة فى المجموعات المرضية والتى تحتاج إلى توجيه وإرشاد، كما يجب أن تستخدم هذه التقارير من المتخصصين ذوى الكفاءة العالية وان تعامل هذه التقارير كوسيلة مساعدة فى التشخيص.

وظهرت فى السنوات الأخيرة أيضًا نظم أكثر تقدما تشمل ما يسمى «بنوك الأسئلة» أو « مخازن مفردات الاختبارات»، والتى تقوم فيها الحاسبات الالكترونية بالدور الحاسم، وتفيد هذه «البنوك» أو «المخازن» فى انتقاء المفردات أو الأسئلة لإعداد اختبار بمواصفات معينة (مثل أهداف تعليمية معينة، محتوى معين، مستوى صعوبة معين... إلخ)، وعادة ما يشمل النظام الكامل لبناء الاختبارات عن طريق الحاسب الالكترونى المكونات الآتية:

١ – بنك الأسئلة.
٢ – نظام لتوليد الأسئلة.
٣ – بنك بخصائص الأسئلة.
٤ – نظام لانتقاء الأسئلة.
٥– الإختبارات التواؤية المحوسبة.

ويعتبر القياس الموضوعى أحد الأسس الهامة التى قامت عليها بنوك الاسئلة، ويقصد ببنك الاسئلة تنظيم مجموعة كبيرة من الاسئلة ثم تقديمها وتدريجها وحفظها بطريقة معينة بحيث يمكن الإضافة لها أو السحب منها، وقد تستخدم ذات الاسئلة مرات متعددة ويمكن استخدامها فى بناء الإختبارات.

ونظرا للزيادة المطردة فى بناء الاسئلة الموضوعية وتطور نظم الحاسبات فإن ذلك أدى إلى امكانية حفظ الأسئلة وتخزينها، وكانت انسب طريقة فى ذلك هو استخدام اشرطة الحاسبات Computer Discs.

ومن خلال بنوك الاسئلة أيضا أمكن تطوير الإمتحانات ونظمها وتوفير الوقت فى بناء الاسئلة بأنواعها، وتدريب المتعلمين على استخدام أنواع متعددة من الاسئلة متدرجة الصعوبة بالإضافة إلى امكانية تحليل استجابات المتعلمين وتشخيص صعوبات الاداء وما قد يترتب على تغيير أساليب التدريس.

ومن المجالات المهمة التى امتد إليها استخدام الحاسب الالكترونى فى التقويم النفسى لبرامج التعليمية التفريدية التى يديرها الحاسب الالكترونى (أى حين يستخدم الكمبيوتر فى تطبيق ما يسمى الاختبارات «التفصيل» tailored أو الاختبارات التكيفية التتابعية، فالمرونة الفائقة التى عليها الحاسب الالكترونى تسمح له بأن يطبق على كل مفحوص الأسئلة أو المفردات الأكثر ملاءمة لمستوى قدرته، وفى هذه الحالة يعتمد الكومبيوتر على استجابة المفحوص على سؤال معين، ويتحدد فى ضوء هذه الاستجابة نوع السؤال أو مجموعة الأسئلة التى تتلوه، ومثل هذا النوع من الاختبارات «التفصيل» توفر الوقت والجهد اللذين يبذلان فى الإجابة على اختبارات شديدة السهولة أو الصعوبة بالنسبة للمفحوص (٣٤أ -٧١٢).

ويوجد استخدام آخر للحاسب الالكترونى يعد أكثر حداثة وهوتزويد الفاحصين ببعض التفسيرات المفيدة فى الإرشاد النفسى ومن ذلك أن يقدم الكمبيوتر تقريرا لفظياً عن أداء المفحوص يصف المهن الأكثر ارتباطا وملائمة لاستعداداته وميوله وتطلعاته التربوية يسمى «تقرير التخطيط المهنى»، وقد لوحظ أن الكمبيوتر يستطيع القيام بعمليات المقارنة بين أنماط درجات الاختبارات المختلفة للمفحوص ويقوم بالربط بينها وبين المتطلبات المهنية المختلفة على نحو أيسر وأسرع، بل وأكفأ، من المرشد النفسى. ومن المتوقع أن يوظف فى المستقبل القريب هذا الجانب من استخدامات الكمبيوتر فى التربية وخاصة ما يتصل بتشخيص صعوبات التعلم المدرسى. وتوجد بوادر لذلك فى الوقت الحاضر بالنسبة لاختبارات المهارات الأساسية.

من العرض السابق لإستخدام الحاسبات فى مجال القياس النفسى والتقويم التربوى فإن هناك مزايا لهذا الإستخدام وأيضا ظهرت بعض العيوب كالآتى:

١- إن الإختبارات تعدل تلقائيا طبقا لمستوى أداء المفحوصين.

٢- تتعدل أيضا بناء على ذلك المحددات السيكومترية للإختبار.

٣- إن زمن الإختبار يتأثر بأداء الفرد، فاستبعاد بعض المفردات طبقا لبرنامج الحاسب فى ضوء اداء المفحوص يصبح قصيرا عما كان محدد حتى إذا كان الإختبار موقوتا.

٤- إن المفحوص المتميز كان يقضى وقتا طويلا فى حل اسئلة سهلة متضمنة فى الإختبار، ومن خلال استبعاد المفردات تلقائيا فى ضوء أداء المفحوص يقل الزمن المستغرق فى الاداء الاختبارى، وكذلك الحال بالنسبة للمفحوصين الضعفاء حين تستبعد ايضا الاسئلة الصعبة والتى يمضون وقتا طويلا فى حلها دون جدوى.

٥- نظرا لأن برنامج الحاسب يستبعد اسئلة فى ضوء الاداء الفعلى للمفحوصين فإنه يصبح ايضا من اليسير إضافة اسئلة طبقا لاداء المفحوص الفعلى.

٦- السرعة الفائقة فى تصحيح أداء المفحوص والحصول على تقارير عن أداءه.

٧- يقلل استخدام الحاسب فى تطبيق وتصحيح الاختبارات التحيز والتمييز بين المفحوص والفاحص.

٨- امكانية الحصول على زمن كل مفرده وزمن كمون الاستجابة، ويمكن من خلال ذلك تحديد صدق المفردات وثباته ومعاملات التمييز.

٩- أما عيوب استخدام الحاسبات فترتبط بالتكلفة فى الإعداد والتنفيذ وكذلك عدد الاجهزة المطلوبة والمتاحة لمثل هذا النوع من الاختبارات (٣٤أ ، ٢٤أ).

الاختبار التواؤمى المحوسب Computerzed Adaptive Testing(*)

لقد تطورت أساليب القياس النفسى فى الآونة الأخيرة تطورا هائلا سواء على المستوى النظرى مثل نظريات القياس المتقدمة، أو على مستوى الممارسة مثل استخدام الحاسب فى تقديم ومعالجة وتحليل المفردات. فعلى المستوى النظرى، ساعد ظهور نظرية الاستجابة للمفردة Item Response Theory (IRT) فى الخمسينيات من القرن الماضى على تطوير أسلوب القياس النفسى فى كافة جوانبه، حيث وضعت اساسا نظريا جديدا للقياس النفسى والتربوى بعد ان سادت نظرية الاختبار التقليدية Classical test theory طوال ما يزيد عن ٥٠ عاما مضت (والتى هى فى الحقيقة

(*) نشكر ا.د. آمال صادق د. أمين صبرى على اضافة هذا الجزء.

مجموعة من الاجراءات التى تعالج بها المفردات مثل ثبات وصدق وتمييز المفردة بالطرق التقليدية) .

لقد أمكن لعلماء القياس النفسى التغلب على الكثير من المشكلات التى عانت منها نظرية الاختبار التقليدية عن طريق عدد من الأساليب السيكومترية الجديدة، والإستفادة من بعض النماذج الرياضية الحديثة مستثمرين فى ذلك التقدم التكنولوجى الحالى فكان من نتاج ذلك ظهور نظرية الاستجابة للمفردة .

أما على المستوى التطبيقى، فقد تحسنت عملية القياس، وصارت أكثر دقة عن ذى قبل، باستناد الكثير من الاختبارات النفسية والتحصيلية على نظرية الاستجابة للمفردة، فقد بنيت العديد من اختبارات التحصيل والاستعداد المقننة على مبادئ تلك النظرية (٣٤١:٣٥)وتمكن السيكومتريون كذلك من بناء بنوك للأسئلة يمكن من خلالها إنشاء صور متكافئة من الاختبارات كما تتيح نظرية الاستجابة للمفردة أيضا إمكانية تشبيك linking عدد من الاختبارات التى تقيس متغيرا واحدا بحيث تصير مقياسا واحدا تتدرج مفرداته على ميزان تدرج واحد مشترك.

وقد واكب هذا التطور المتلاحق للقياس النفسى طفرة هائلة فى تقنيات الحاسبات سواء فى مجال معدات الحاسبات Hardware أو البرمجيات Software خاصة مع بداية الثمانينيات. وكان لابد للقياس النفسى والتربوى ان يساير هذا التطور، ويستثمر تلك التقنيات لتحسين اجراءات القياس. وكان من جراء استخدام القياس النفسى للتطرات التقنية الحديثة ان ظهرت اساليب جديدة للاختبار على الحاسب احتلت مكان اختبارات الورقة والقلم التقليدية لتطبيق الاختبار فى سياقات عديدة (١٨٨:٤٠) .

وقد أدى كل ذلك إلى ظهور مصطلحات عديدة لتعبر عن أشكال مختلفة من القياس وعلى رأسها مصطلح الاختبار المحوسب Computerized Testing ويتم تطبيق الاختبارات المحوسبة إما عن طريق تقديم جميع مفردات المقياس معروضة على الحاسب لجميع المفحوصين وبنفس الترتيب (٢٠-١٩٩٧) ويسمى ذلك النمط بـ «لاختبار الخطى المحوسب» ، أو يقدم ذلك الاختبار المحوسب بطريقة تواؤمية adap-tive حيث لا تقدم للمفحوص إلا المفردات التى تتناسب مع مستواه فقط، وهو ما

يسمى بت «الاختبار التواؤمى المحوسب» Computerized Adaptive Testing (CAT) وفى ذلك تعرض على المفحوص من خلال الحاسب بعض المفردات لتحديد مستواه المبدئى، وبناء على أدائه على تلك المفردات الاستهلالية تقدم له مفردات الاختبار التى تتناسب مع مستواه. ويعتمد اختيار مفردة ما على سلسلة استجاباته للمفردات السابقة، وخصائص تلك المفردات (١٢:٤٠) ونظرا لأن الحاسب يعيد تقدير قدرة المفحوص بعد كل مفردة، وينتقى له المفردات الأكثر ملاءمة، فإنه من الممكن الوصول إلى مستويات عالية من الدقة لتقدير قدرة الفرد باستخدام عدد محدود من مفردات الاختبار. وقد أصبح بذلك مفهوم «الاختبار التواؤمى المحوسب» (CAT) هو أهم وأحدث تطبيقات نظرية الاستجابة للمفردة فى مجال القياس النفسى الحديث (٣٨-١٩٩٩).

ولا يشترط فى الاختبارات التواؤمية المحوسبة الإجابة على جميع الاسئلة أو بنفس الترتيب، إنما يختلف عدد المفردات المطبقة وترتيبها من فرد لآخر، حسب قدرة كل فرد.

وللاختبار التواؤمى المحوسب أركان محددة هى: اختيار نموذج استجابة للمفردة، وبناء مستودع المفردات، وأسلوب انتقاء المفردة، وأسلوب تقدير القدرة، وأسلوب قاعدة الإنهاء.

وقد اجريت مئات البحوث حول الاختبارات التواؤمية المحوسبة من مختلف جوانبها والنتيجة العامة لتلك الدراسات هى ان للاختبار التواؤمى المحوسب معاملات ثبات وصدق مكافئة لمعاملات ثبات وصدق الاختبارات التقليدية أو اعلى منها حتى مع خفض طول الاختبارات إلى حوالى ٥٠٪ ، بالرغم أنها تقيس نفس المتغيرات (٣٧-١٩٧٧) وتراوح انخفاض زمن لاختبار من ٢٥٪ إلى ٧٥٪ من زمن الاختبار الاصلى دون فقدان دقة القياس. كما تم الوصول إلى تقديرات ثبات بلغت ٠,٨٠ عن طريق تطبيق ٩ مفردات فقط (٢١ :٢٠٥).

وهناك عدد من المؤسسات والهيئات فى الولايات المتحدة الأمريكية والدول الغربية اعتمدت الاختبار التواؤمى المحوسب فى انظمتها مثل مؤسسة الدفاع الأمريكية التى تبنت إصدار بطارية الاستعداد المهنى للقوات المسلحة (٣٩ :١٩٩٧) وكذلك

مؤسسة الاختبارات التربوية التى حولت عددا من الاختبارات الشهيرة إلى اختبارات تواؤمية محوسبة مثل اختبار اللغة الانجليزية كلغة اجنبية (TOEFL) واختبارات (GRE) وغيرها من الاختبارات. كما تبنت شركة مايكروسوفت الشهيرة فى نظم الحاسبات والبرمجيات الاختبارات التواؤمية المحوسبة فى اختباراتها لنظم التشغيل وبناء البرامج والشبكات (١٨٧:٣٨) كما أجريت عدد من الدراسات أيضا فى مجال اختبارات الشخصية التواؤمية المحوسبة (٣١-٢٠٠٥).

بعض المشكلات المعاصرة فى ميدان التقويم النفسى:

يواجه ميدان التقويم النفسى فى الوقت الحاضر عدداً من المشكلات تتصل فى جوهرها باستخدام الاختبارات، ويمكن أن نلخص هذه المشكلات فيما يلى:

١ - مستوى جودة الاختبارات.

شهد مطلع الستينيات بداية حملة منظمة ضد الاختبارات المقننة سرعان ما اشتدت وزاد سعيرها فى السبعينات ثم الثمانينيات من القرن الماضى، وقد واجهت سهام النقد على وجه الخصوص إلى الأسئلة الموضوعية من نوع الإختيار من متعدد التى تتألف منها معظم هذه الاختبارات، واتهمت بأنها تعاقب المفحوص الأكثر ذكاء وابتكارية، وقد دعم النقاد هذا الرأى بما يلاحظ من أن هذا المفحوص الذكى المبتكر يظهر فى إجابته على أسئلة الاختبارات المقننة اختلافا عن الإجابات التى تتحدد صحتها مقدما والتى تتكون منها مفاتيح هذه الاختبارات، وعادة ما يحكم على هذا النوع من الإجابات التى لا تتفق مع المفتاح بالخطأ بالرغم من أننا لو حللناها بعمق وأناة نجد فيها قدراً كبيرا من الصواب، وهذا النقد بالطبع ما يبرره وخاصة أنه وجدت فى بعض الاختبارات المقننة ذات الشهرة الكلاسيكية أسئلة من هذا النوع. ناهيك بالطبع عما يمكن أن تجده فيما يتوافر لدينا.

إلا أن هذا النقد لا يقلل من شأن الاختبارات الجيدة الصنع المتقنة البناء، وقد كانت له تضميناته الإيجابية، فأصبح ناشرو الاختبارات فى وقتنا الحاضر يقدمون لمستخدميها بيانات تحليلية تفصيلية عن الأسئلة التى يتألف منها الاختبار مع تحليل إحصائى ومنطقى لها. كما أن هذا النقد جعل مستخدمى الاختبارات أكثر حرصاً وحذراً فى استخدامها بالإضافة إلى شعورهم بالحاجة إلى عدم الاعتماد على نتائج

الاختبارات وحدها، وإنما لابد من تدعيمها بمعلومات إضافية من المصادر الأخرى.

ويرتبط بهذا النقد نقد آخر يتصل بمدى شمول الإختبارات لنطاق واسع من السلوك الإنسانى، إننا نعلم أن الاختبارات لا تقيس فى معظم الحالات إلا عينات من السلوك الإنسانى (إذا تجاوزنا الاختبارات المرجعة إلى المحك ذات الاستخدامات الخاصة فى مجال التعليم والتدريب)، ومع ذلك نقع فى الخطأ الشائع الذى يتمثل فى أننا نتوقع منها ما لم تقصد إليه، ولعلنا نكون أكثر علمية ودقة إذا استخدمنا الاختبارات فى تحقيق الوظائف المحددة التى تسعى إليها، أنها حينئذ تصبح أدوات جيدة للعلم، ولعل هذا القول يعود بالنقاد الذين يوجهون إلى اختبارات الاستعداد ومنها اختبارات الذكاء فى وقتنا الحاضر عاصفة من الهجوم الذى جانبه الصواب. إن اختبار الاستعداد الذى يستخدم فى أغراض الانتقاء لمؤسسة تعليمية أو مهنية مثلا لايحقق الغرض منه إذا افترض فيه كل من الناقد والمدافع جميعا أنه يقيس جميع المتطلبات اللازمة للنجاح فى دراسة معينة أو مهنة معينة، فهذا القول يتجاوز المعطيات والبيانات التى توفرها الاختبارات، وحتى لا نقع فى أى تطرف جديد علينا دائما أن نتذكر أن الدرجات التى نحصل عليها من الاختبارات ليست إلا نوعا من المعلومات ـ من بين أنواع عديدة أخرى متاحة للأخصائى النفسى، ولايجب عليه أن يصدر حكمه أويتخذ قراره بناء على نتائج الاختبارات وحدها.

إن الاختبارات النفسية ـ كغيرها من وسائل أدوات البحث ـ قد تتفاوت فى درجة الدقة، وعلينا ألا نتوقع منها ما ليس من طبيعتها، وعلينا أن نجعل نصب أعيننا دائما أنه إذا كانت المقاييس الفيزيائية ـ وما هى عليه من سمعة وتاريخ ـ تعانى من أخطاء القياس وأخطاء الملاحظة وأخطاء التنبؤ، فما بالك بالاختبارات النفسية وهى مقاييس للسلوك الإنسانى بكل ما فيه من تعقد وتغير ورقى.

٢ – أثر الاختبارات فى المفحوص:

يوجه النقاد إلى الاختبارات نقدا آخر يتلخص فيما تحدثه هذه الاختبارات من آثار غير مرغوب فيها ـ إن لم تكن ضارة ـ فى المفحوصين من مختلف المستويات العمرية، ولعل أشهر هذه الآثار ذكرا هو القلق. وبالطبع لاينكر أحد أن الموقف الاختبارى يحدث فى المفحوص قدراً من القلق، إلا أنه من الحقائق السيكولوجية

المهمة أن قدراً معقولا من القلق يدفع المفحوصين إلى مزيد من التحسن فى الأداء، وقد يكون الأثر الضار للقلق مرتبطا بالمفحوصين ذوى المستويات العليا فيه حيث يتداخل القلق الزائد مع الأداء الاختبارى، وقد لايكون فى ذلك مسئولية مباشرة للاختبار ذاته، فكل ما يفعله القلق مع هؤلاء المفحوصين ذوى القلق المرتفع أن يضيف إليه المزيد، ويوجد فى فن تطبيق الاختبارات قواعد أساسية لو التزم بها الفاحصون يمكنهم خفض القلق الاختبارى، ويمكن أن يتم ذلك فى جميع مراحل تطبيق الاختبار، سواء عند التهيؤ له، أو التدريب على الأمثلة التوضيحية أو استخدام الوقت الحر فى الأداء الاختبارى (اختبارات القوة مثلا) . ولحسن الحظ يتحول ناشرو الاختبارات فى الوقت الحاضر إلى تزويد المفحوصين باختبارات يمكنهم التدريب عليها وتتوافر فى مكتبات الجامعات الأمريكية والأوربية أمثلة كثيرة من هذه الاختبارات التدريبية ، وبالإضافة إلى ذلك فإن مؤلفى اختبارات الأداء الأقصى يتحولون أيضاً فى الوقت الحاضر إلى التركيز على اختبارات القوة (التى لا تتحدد بوقت معين) بدلا من اختبارات السرعة، وفى جميع الأحوال لابد للفاحص من ملاحظة المفحوص بعمق أثناء الموقف الاختبارى، وإعادة النظر فى درجات المفحوصين ذوى الانفعالية العالية وخاصة أن قلق الاختبار ليس له إلا أثر مؤقت فى الصحة النفسية للمفحوصين، ومرة أخرى فإن المسألة تعتمد على الاستخدام الجيد والتوظيف الرشيد للاختبار النفسى ونتائجه .

٣ - الأثر الاجتماعى والثقافى للاختبارات النفسية :

لعل هذه المسائل هى أخطر المسائل على الإطلاق، وتأتى أهميتها من أنها شغلت الباحثين فى مجال التقويم النفسى وخارجه سنوات طويلة تكاد تتساوى مع العمر الزمنى لحركة القياس النفسى ذاتها. فكثيرا ما يقال أن الاختبارات النفسية غير عادلة فى تعاملها مع الفئات الاجتماعية داخل المجتمع ومع السلالات البشرية بين المجتمعات المختلفة . وقد تزايدت هذه الانتقادات فى السنوات الأخيرة فى بعض المجتمعات التى مورست فيها الاختبارات النفسية لسنوات طويلة، وعلى الأخص بريطانيا والولايات المتحدة الأمريكية، ويرجع ذلك إلى زيادة المطالبة بالحقوق المدنية وتهيئة تكافؤ الفرص للجميع.

وتتلخص الاتهامات الاجتماعية التى يوجهها النقاد إلى الاختبارات النفسية بأنها متحيزة اجتماعيا وتمثل حواجز ضد الفرص التعليمية والمهنية لعدد ن أبناء الفئات المحرومة داخل المجتمع الواحد.

وحتى يمكن لنا مناقشة هذه الاتهامات بشيء من الموضوعية نذكر أن مسألة التحيز أو عدم العدالة فى الاختبار تشير إلى مسألتين مختلفتين، أولاهما تتصل ببناء الاختبار ذاته حيث يتألف من أسئلة أو مفردات تلائم فئات اجتماعية دون سواها، وثانيتهما تتصل بالاستخدام الظالم لنتائج الاختبارات وخاصة فيما يتصل بإصدار الأحكام أو اتخاذ القرارات ، وهى مسألة تخرج بالباحث من نطاق العلم إلى التحيز الصريح.

أما فيما يخص المسألة الأولى فإننا نجد أن معظم ما يكتب عن تحيز محتوى الاختبار يتركز على أداء بعض الجماعات أو الأفراد فى الاختبار يكون متخلفا ، فيحصلون على درجات منخفضة بالمقارنة بغيرهم من الجماعات أو الأفراد، إلا أن هذا ليس حجة ويتجاوز الدليل الموضوعى، فالأداء المنخفض فى ذاته ليس برهانا كافيا على تحيز الاختبار، وإلا كان كل اختبار فى الحساب يصبح متحيزا ضد الضعاف فى الرياضيات، وكل اختبار فى المفردات اللغوية يصبح متحيزا ضد أصحاب المحصول اللغوى المتواضع، إن هذا القول لا يصبح صحيحا إلا إذا تدخلت فى الاختبار عوامل أخرى لم يتحكم فيها الباحث، كأن تصاغ المشكلات الحسابية فى قالب لغوى يحتاج إلى قدرة لغوية عالية، أما إذا تحرر الاختبار الحسابى من هذه العوامل (أى كان صادقا بالمعنى الدقيق للكلمة) فلماذا لايكون ضعف الأداء فيه دليلا على نقص فى القدرات والمهارات الحسابية تحتاج إلى نوع من التربية التعويضية أو التعلم العلاجى.

إلا أن ما نحب أن ننبه إليه هو أن بعض الاختبارات المقننة تركز على المحتوى والقيم الثقافية المألوفة والشائعة لدى الإنسان الغربى الأبيض من الطبقة المتوسطة (فقد كانت الاختبارات منذ نشأتها الأولى غربية المنشأ وفى قالب لغوى يحتاج إلى قدرة لغوية عالية، أما إذا تحرر الاختبار الحسابى من هذا العامل أصبح ملائما، ولذلك لابد من أن يتنبه الباحثون فى الثقافات الأخرى إلى عدم اللجوء إلى «النقل الأعمى» أو

«الترجمة الحرفية» لهذه الاختبارات، وإلا أوقعوا أنفسهم وثقافتهم فى مأزق اجتماعى خطير، ولعلنا بهذا نشير إلى بعض الممارسات الخاطئة فى حركات الاختبارات النفسية فى الثقافتين المصرية والعربية حيث ينطبق على معظمها خصائص النقل الأعمى والترجمة الحرفية، بل أننا بهذا ندعو إلى إبداع ثقافى جديد فى مجال الاختبارات النفسية، فليس كل اختبار ظهر فى عاصمة أوربية أو فى جامعة أمريكية قابل للنقل، فبعضه وراءه من الافتراضات النظرية فى علم النفس والأسس الثقافية للمجتمع ما يجعل نقله - حتى ولو كان مع بعض التعديل - نوعا من «المجزرة» العقلية كما قال - بحق - عالم النفس التركى جوندوز فاسف فى بحثه عن الاختبارات النفسية فى العالم الثالث والذى ألقاه فى المؤتمر الدولى الثالث والعشرين لعلم النفس الذى انعقد بالمكسيك عام ١٩٨٤ -

دور التقويم فى التحقق من جودة التعليم*

إن المجتمع المصرى شأنه فى ذلك شأن المجتمعات المعاصرة – يتعرض لتغيرات اقتصادية وسياسية واجتماعية وثقافية متنامية ومتسارعة نتيجة عوامل متعددة من بينها التقدم العلمى والتكنولوجى وثورة المعلومات والإتصالات، ومن المتوقع أن هذه التغيرات تزداد حدة وتسارعا.

ولعل أهم التغيرات التى شهدتها مصر فى نهاية العقدين الأخيرين من القرن العشرين ما يلى:

١ – التحول الاقتصادى من الاقتصاد المخطط إلى الاقتصاد الحر.

٢ – ظهور التعددية السياسية من خلال نظام حزبى يسمح بالتعبير عن مختلف الآراء.

٣ – إنشاء مؤسسات يمكن أن يكون لها دور أكثر فعالية فى عمليات الديموقراطية والمشاركة.

---

* بتصرف عن الورقة التى أعدها فؤاد أبو حطب إلى المجلس القومى للتعليم ١٩٩٩/١٠/١٦ وموضوعها «الإرتقاء بمستوى خريج التعليم الجامعى والعالى فى اطار مفهوم الجودة الكلية لمواجهة تحديات المستقبل» – منشورة فى المجلة المصرية للدراسات النفسية العدد التذكارى رقم ٣١ ابريل سنة ٢٠٠١ .

٤- تنفيذ مشروعات للتنمية الشاملة والمستمرة وخاصة المشروعات الكبرى التى تسعى بالخروج من دائرة الوادى الضيق إلى نطاق أرض الوطن الواسعة وبالطبع فإن مصر لاتزال تواجه تحديات ومشكلات لم تحل على مدى القرن الماضى ويتطلب ذلك التركيز على القضايا الأساسية المرتبطة بتوسيع نطاق الديموقراطية المتاح على نحو يسمح بمشورة أكثر إيجابية وفعالية، كما يتطلب أيضا النهوض بالمستوى الاقتصادى والثقافى والمعرفى والبشرى على نحو يسمح بالمنافسة على المستوى الاقليمى والعالمى كما يتطلب الأمر أيضا رفع مستوى البحث العلمى وتطويره ودعمه وربطه بمشروعات إنتاج المعرفة والتكنولوجيا، كما أن التصدى لسلبيات التعليم بكافة مستوياته فى المراحل التعليمية المختلفة، والعمل على بناء منظومة تقويم له من منظور الجودة الكلية أصبح من الضرورات العاجلة.

وعلى الرغم من الإسهامات الكبيرة للتعليم فى بناء مصر الحديثة والتضحيات التى قدمتها الدولة لتطوير التعليم والإرتقاء بمستواه، إلا أن هناك شعوراً عاماً بأن التعليم يعانى من مشكلات عديدة وأوجه قصور شتى من أهمها تدنى مستوى مخرجاته وقد يكون أكثرها وضوحاً ضعف مستوى الخريجين سواء فى التعليم قبل الجامعى أو التعليم الجامعى والعالى.

وقد صاحب هذا الهبوط فى مستوى خريجى التعليم الجامعى والعالى التى - يشعر بها الجميع - بضع مشكلات لعل أخطرها ما يسمى مرض الشهادة Diploma Disease" والتى يسعى الطلاب للحصول عليها بأية وسيلة، وقد ترتب على ذلك تدنى أساليب التدريس وتخلف المناهج، وقصور البيئة التعليمية، وتفشى الأمراض التربوية الشائعة فى منظومة التعليم كلها وقد يكون أخطرها الدروس الخصوصية.

ولعل أخطر الآثار السلبية التى ترتبت على ذلك عدم الاعتراف ببعض الشهادات بصفة خاصة فى المستوى الجامعى، وتدنى مستوى الخريجين إذا ما قورنوا بالمستويات العالمية.

وبالطبع لا يمكن للأحوال الراهنة لمستوى الخريجين أن ترضى أحداً أو تقنعه بأن هذه النوعية من الخريجين يمكن أن تقود المجتمع إلى عالم القرن الواحد والعشرين بتحدياته المختلفة.

وحتى تتهيأ المؤسسات التعليمية فى مصر (سواء على المستوى قبل الجامعى أو الجامعى والعالى) لتحمل أعباء المستقبل، لابد لها أن تسعى إلى تحقيق الجودة الكلية فى مكونات عملية التعليم بحيث تؤلف مكوناتها منظومة شاملة، وهذه المكونات هى:

أ – المدخلات وتشمل

أ–١ مدخلات الطلاب بما فى ذلك نظام القبول فى التعليم.

أ–٢ مدخلات المعلمين وإختيارهم وتنميتهم وتدريبهم.

أ–٣ خصائص البيئة العامة المحيطة بالمؤسسة التربوية أو المؤسسة التربوية ذاتها.

ب– العمليات: وتشمل

عمليات التعليم والتعلم وتشمل:

ب–١ – البرامج والمناهج والمقررات.

ب–٢– الكتب وبرامج التعليم وتكنولوجيا التعليم.

ب–٣– الأجهزة وتجهيزات المعامل والورش.

ب–٤– المكتبة ومراكز تكنولوجيا التعليم

ب–٥– طرق التدريس والتعلم.

ب–٦– رعاية ذوى الاحتياجات الخاصة.

ج– عمليات التقويم

ج–١– التقويم المبدئى وطرق التعليم التعويضى.

ج–٢– التقويم التكوينى وطرق التعليم العلاجى والإثرائى.

ج–٣– التقويم التجميعى للإنتقال من مستوى تعليمى إلى مستوى تعليمى آخر أو من فرقة دراسية إلى فرقة دراسية أخرى فى ذات المرحلة التعليمية.

د– المخرجات وتشمل:

د–١– النواتج التعليمية المقيسة (باستخدام الامتحانات وأدوات التقويم).

د-٢- النواتج التعليمية غير المقيسة (معرفية - أخلاقية - اجتماعية - شخصية).

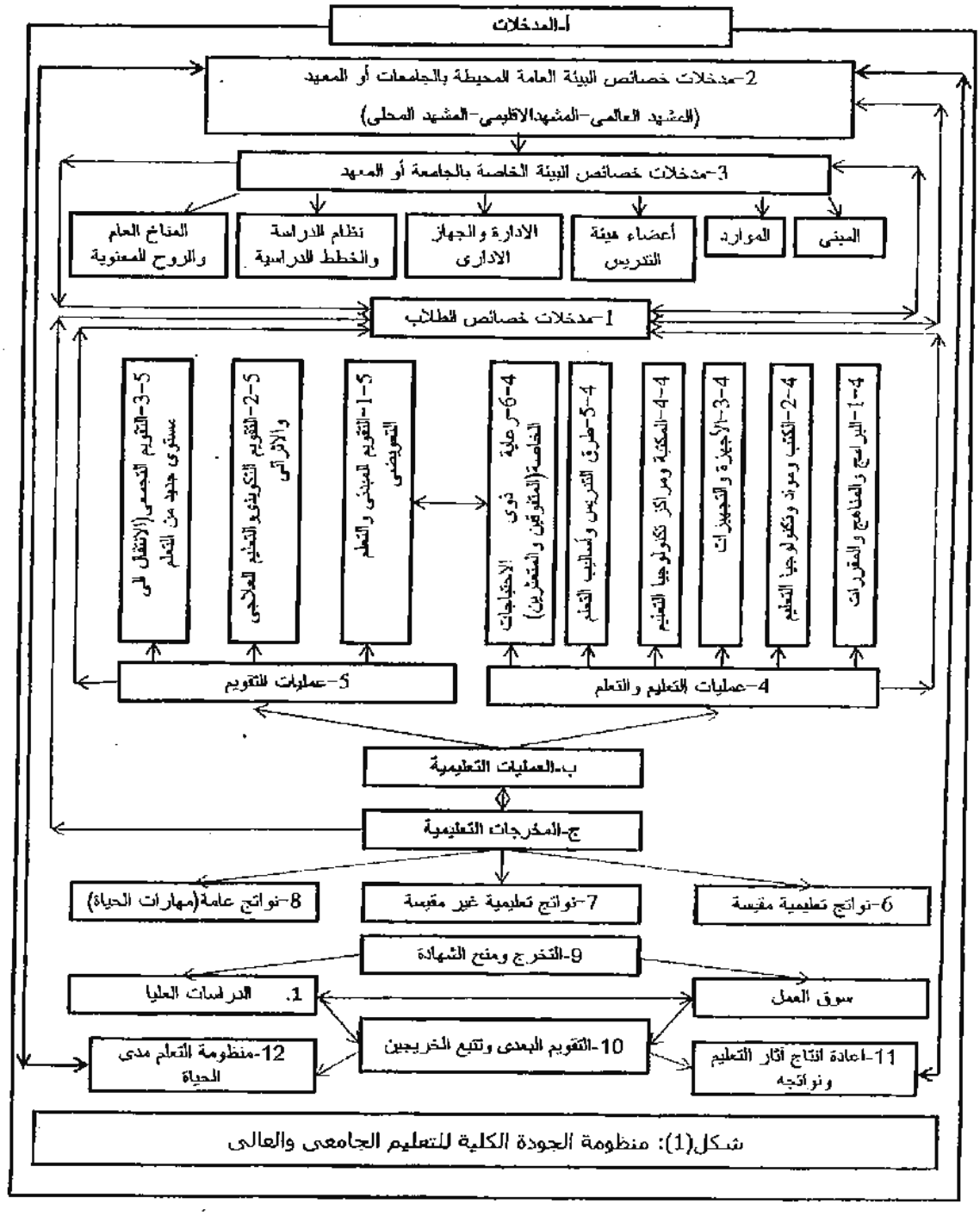
د-٣- نواتج عامة أو مهارات حياتية (إقتصادية - ثقافية - مهنية).

د-٤- التخرج ومنح الشهادة.

د-٥- التقويم البعدى (تتبع الخريجين).

د-٦- إعادة إنتاج أثار التعليم ونواتجه فى الأجيال التالية.

د-٧- منظومة التعليم مدى الحياة.



شكل(1): منظومة الجودة الكلية للتعليم الجامعى والعالى

عن فؤاد أبو حطب ١٩٩٩ .

آلية الجودة الكلية فى التعليم قبل الجامعى والجامعى والعالى

تعد جودة التعليم فى المراحل التعليمية المختلفة مطلبا أساسى لرفع كفاءة مخرجات العملية التعليمية فى المراحل التعليمية المختلفة بما يعكس أثره على مستوى الخريجين ورفع كفاءتهم والإرتقاء بمستوياتهم، بحيث تصبح الشهادات التى يحصلون عليها معترف بها عالمياً، وهو ما يفتقده الخريج الحالى فى كثير من الأحيان.

وقد أخذت معظم الدول بمفهوم الجودة الكلية، وأنشأت آليات لتحقيقها فى مختلف مراحل التعليم، ولعل من أشهرها آليتان هما: آلية الاعتماد أو الإجازة Accredita-tion السائد فى نظام التعليم فى الولايات المتحدة الأمريكية، وثانيهما آلية ضمان الجودة Quality Assurance المعمول به فى نظم التعليم الأوروبية وخاصة فى المملكة المتحدة (فؤاد أبوحطب ١٩٩٩) ونحن فى مصر فى حاجة إلى إستحداث آلية للجودة الكلية لجميع مراحل التعليم، مستفيدين فى ذلك من خبرات الدول المتقدمة وتتألف هذه الآلية من ثلاثة عناصر أساسية هى:

أ- تحديد المرجعيات:

المعايير القياسية أو المستويات والتى يجب توافرها فى كل مكونات المنظومة بحيث تحوز رضا المستفيدين منهم (مؤسسات المجتمع) ورضا المؤسسة المنتجة لهم عن أدائها (المدرسة – المعهد – الجامعة)، ورضا الخريجين عن أنفسهم من خلال تحقيق أهداف الإتقان والتميز ويحتاج تحديد المرجعيات إلى أعلى مستويات الخبرة تكون على وعى بالمشاهد العالمية والإقليمية والمحلية، وتتسم بالمرونة للإستجابة للتغيرات التى تطرأ على هذه المرجعيات، وأن يحدث فيها التطوير المنشود بالسرعة والكفاءة الواجبة.

ب- ضمان ومراقبة الجودة الكلية بالتعليم (قبل الجامعى والجامعى والعالى)

يحتاج مفهوم الجودة الكلية فى التعليم إلى وجود آلية للتقويم الداخلى للجودة، تكون مسئولة عن ضمان الجودة ومراقبتها أو التحكم فيها، وهذه الآلية تقوم بالمراجعة المستمرة والتحقق الدائم للتأكد من ان المخرجات التعليمية (وأى عملية مرتبطة بإنتاجها) تتوافر فيها المواصفات التى تحددها المرجعيات (أو المستويات) المحددة لها، مع اقتراح وسائل الإرتقاء بعناصر المنظومة التعليمية للنهوض بمستوى المنتج البشرى للتعليم.

ومعنى ذلك أن آلية التقويم الداخلى للتعليم قبل الجامعى والجامعى العالى يجب ان يشمل ما يضمن متابعة المرجعيات لكل مكون من مكونات منظومة الجودة .:

التقويم النفسى ـــــــ

٥- التقويم كمنظومة تغذية راجعة:

التقويم فى منظومة التحكم الذاتى التى تعتمد على مسلمة التفاعل المتبادل بين المكونات يتجاوز بكثير حدود محض إتخاذ قرار بالنجاح أو الفشل، الجودة أو الرداءة. فإذا كان المعنى المباشر للتقويم فى التربية والتعليم والتدريب هو الحكم على مدى تحقيق الأهداف بإعتبارها التغيرات والتعديلات المتوقعة فى سلوك الدارسين والمتعلمين فإن ذلك يجب أن يتضمن ما يلى:

١- ملاحظة وقياس وتسجيل مقدار ما يتحقق من الأهداف، أى التغيرات التى تطرأ على سلوك الدارسين بالفعل نتيجة لتعرضهم للبرنامج، ونسمى هذه التغيرات الحادثة والمتحققة نواتج التعلم Learning Outcomes وتتطلب ملاحظة وقياس وتسجيل نواتج التعلم هذه إستخدام الأدوات المناسبة (الاختبارات وغيرها من طرق جمع المعلومات).

٢- حساسية للظروف والأحوال المختلفة التى يتعرض لها تنفيذ البرنامج من خلال المواد الدراسية والوسائل التعليمية وطرق التعليم أو التدريب وأساليبه المختلفة، سواء من داخل منظومة البرنامج أو من خارجها (التشويش noise).

٣- مقارنة مستمرة بين الأهداف كتغيرات متوقعة فى سلوك الدارسين وما يتحقق منها الفعل فى صورة نواتج تعلم.

٤- الحكم على مدى تحقيق الأهداف فى ضوء المقارنة السابقة.

٥- اتخاذ قرار على ضوء هذا الحكم بإنتقال الدارس إلى مستوى جديد من التعلم حين تتحقق الأهداف، أما إذا لم تتحقق الأهداف فإن معلومات التغذية الراجعة التى يوفرها التقويم فى هذه الحالة تمتد إلى جميع مكونات المنظومة.

٦- دراسة الآثار التى تنشأ عن العوامل أو الظروف التى تسهل الوصول إلى الأهداف أو تعطله بإستخدام معلومات التغذية الراجعة المشار إليها. فقد نلاحظ أن هناك قصورا فى تحقيق الأهداف بشكل أو آخر، أو أنه لم يطرأ على سلوك الدارسين تغير يذكر على الرغم مما تعرضوا له من عمليات تعليم أو تدريب. وقد نكتشف نتيجة لذلك مثلا أن هناك نقصا أو قصورا فى هذه العمليات أو ان الاهداف كانت أعلى أو أدنى كثيرا مما يجب، أو أدوات التقويم التى استخدمناها لم تكن جيدة ولا تتوافر فيها الخصائص الواجبة.

وهكذا فإن معلومات التغذية الراجعة التى نحصل عليها من التقويم تجعلنا نعيد النظر فى بعض أو معظم أو جميع مكونات منظومة التعليم أو التدريب (ومنها مكون التقويم ذاته) وذلك تعديلا للمسار وتصحيحه على نحو يجعل المنظومة محققة بالفعل لاهدافها.

وهكذا يصبح التقويم فى هذا التصور أكثر تعقيدا من أن يكون محض إصدار أو خاتمة مطاف كما هو الحال فى التصور التقليدى الخطى لمنظومة التعليم. ويصبح له موضع فى جميع مراحل هذه المنظومة كما نتصورها كمنظومة تحكم ذاتى (١٢ :٦٥١) .

دور التقويم فى تحديث التعليم فى مصر

قلنا أن العملية التعليمية يمكن وصفها بأنها منظومة أو نسق يتألف من حلقات مرتبطة متفاعلة يؤثر بعضها فى بعض ويتأثر به. ويمكن لعملية التفاعل هذه أن تتم بنجاح لو كانت جميع حلقات هذه المنظومة أقل مقاومة للتغيير والتطوير، إلا أن ما يحدث عادة ان بعض هذه الحلقات يقاوم التغيير وقد يكون التقويم أقوى هذه الحلقات مقاومة للتطوير، ومع ذلك فإنه يكاد يكون أكثر الحلقات أثرا فى المنظومة كلها فمن المعلوم ان طبيعة الامتحانات (وهى من وسائل التقويم التربوى) تحدد مسار عملية التعلم ووجهتها.

فمهما تحدثنا عن الأهداف العليا للتربية وعن تطوير المناهج وتحسين التدريس، فإن الامتحانات تحدد لنا فى النهاية ما يتم تعلمه بالفعل. وتكاد تكون صورة لكل ما تم بالفعل، كما تحدد مسار ما يتم. فمثلا نستطيع أن نكتب أو نتحدث كما نشاء عن الفهم والابتكار والقيم وغيرها من الأهداف «السامية» للتعليم وقد تطورت المناهج والكتب الدراسية والوسائل التعليمية فى هذا الاتجاه.

وقد تدرب المعلمين على ذلك وقد يسعى المعلمون إلى تحقيق هذه الاهداف بالفعل. ثم تأتى الامتحانات فتقيس الحفظ والاستظهار وحدهما، وفى هذه الحالة نجد أن كل حديثنا عن التفكير والإبداع وغيرهما لا يتعدى حدود «التلفظ» التربوى إلى الممارسة التربوية الفعلية، فالتلميذ سوف يحفظ والمعلم سوف يحفظ تلاميذه لان الامتحانات تقيس الحفظ وحده.

وهكذا تلعب الإمتحانات المختلفة دورا خطيرا فى تخلف التعليم بل نكاد نقول أن الامتحانات هى مرآة النظام التعليمى كله، وهى مثل مرآة «سنووايت» تعكس الحقيقة لا المظهر، بل نكاد نصل بقولنا هذا إلى حد القول : إذا أردت أن تدرس نظاما تعليميا بفلسفته وقيمته وأسسه وأصوله وأهدافه وأساليبه وممارسته ونواتجه أدرس امتحاناته أو بعبارة أكثر شمولا : إدرس نظم التقويم فيه.

ولذلك فإن دعوتنا إلى تطوير الامتحانات وأساليب التقويم فى مصر وتحديثها تتضمن فى جوهرها تطويرا وتحديثا للمنظومة التعليمية كلها حيث يؤدى إلى تطوير الاهداف التعليمية تحديدا وصياغة، وتطوير إعداد المعلمين قبل الخدمة وتدريبهم أثناءها وتطوير المناهج والمقررات (المحتوى) وتحسين طرق التدريس واساليبه ووسائله، والكتب المدرسية وحل بعض المشكلات التى يواجهها التعليم المصرى فى الوقت الحاضر مثل كتب الملخصات والدروس الخصوصية والغش فى الامتحانات وفى النهاية تغيير البناء النفسى للمتعلم حيث يصبح مواطنا تتوافر فيه الخصائص التى تتلاءم مع متطلبات المجتمع والعصر وتوقعات المستقبل وتحتل فيه مهارات الإنتاج وقدرات المرونة والاستدلال والإبداع محل صناعة الكلام والحفظ والاستظهار والاسترجاع (١٢: ٦٥٥-٦٥٧)

٧- معالم لتحديث التقويم فى مصر:

يمكن أن نحدد المعالم الاتية لأى محاولة لتطوير التقويم فى النظام التعليمى فى مصر وتحديثه

١- الخطوة الأولى لكل من التعليم والتقويم هى تحديد أهداف التعليم تحديدا واضحا قابلا للملاحظة، كما يتمثل فى نواتج التعلم المتوقعة وتتوافر فى الوقت الحاضر طرق متنوعة لتحديد الأهداف وصياغتها على نحو يجعلها قابلة للملاحظة بحيث توجه هذه الصيغ عمليات التعليم والتقويم وتفيدنا فى هذا الصدد نتائج البحث فى ميدان القدرات العقلية.

٢- ليست اختبارات الورقة والقلم (الإمتحانات التحريرية) هى الأسلوب الوحيد للتقويم، كما هو شائع الآن، بل ليست الأسلوب الأمثل فى بعض المجالات بالضرورة. فمن المعلوم أننا فى المادة الدراسية الواحدة نحتاج لأكثر من وسيلة تقويم تبعا لطبيعة المادة وخصائص الهدف التعليمى الذى نسعى اليه من خلال تدريسها والمستوى الذى تقدم فيه هذه المادة. ولذلك فنحن فى حاجة إلى تنوع اساليب التقويم فى ضوء هذه الاعتبارات الثلاثة، ما بين إختبارات مقال وإختبارات تتطلب إجابات قصيرة، واختبارات موضوعية بأنواعها المختلفة، ومقاييس تقدير وقوائم ملاحظة تستوعب ما تتطلبه المادة (كما تحدده أهداف تدريسها) مع أداء تحريرى أو شفوى أو عملى أو غير ذلك.

٣- إن كل مالا يخضع للتقويم الفعلى والفعال والجاد من الأهداف التعليمية يؤدى إلى إهماله فى التدريس الفعلى. فحيثما لا تهتم الامتحانات بالتقويم الجاد بطرق

التقويم النفسى ـــــــ

الاساسية للعملية التعليمية. وقد قام الراحل العظيم عالم النفس المصرى فؤاد أبو حطب بإنشاء المركز ورئاسته منذ انشاءه وحتى عام ١٩٩٨ .

والحقيقة ان وزارة التربية والتعليم قد أدركت منذ وقت مبكر أهمية دور الامتحانات والتقويم التربوى فى اصلاح التعليم فقد سبق انشاء المركز خطوات جادة على الطريق تمثلت فى:

١- إنشاء المجلس الاعلى للامتحانات والتقويم التربوى فى ١٩٨٧/١١/٢٩ وقد أصدرت الامانة العامة لهذا المجلس نماذج لاسئلة امتحانات نهاية الحلقة الابتدائية ونهاية مرحلة التعليم الاساسى وشهادة الثانوية العامة فى اعوام ١٩٨٨/١٩٨٩، ٨٩/٩٠، ٩٠/٩١ بالاضافة الى مشروع انشاء بنوك الاسئلة.

٢- إنشاء مركز تطوير الامتحانات والتقييم التربوى بوزارة التربية والتعليم بالقرار الوزارى رقم ١٨٨ فى ١٩٩٠/٨/١٢ وفيه أصبح المجلس الأعلى للامتحانات والتقويم التربوى هو السلطة العليا المهيمنة على شئون هذا المركز.

وبصدور القرار الجمهورية رقم ٤٦٢ فى ١٩٩٠/١١/١٨ بانشاء المركز القومى للامتحانات والتقويم التربوى كهيئة عامة لها الشخصية الاعتبارية ومن الموسسات العلمية فى الدولة حل المركز الجديد محل جميع المؤسسات السابقة ومنها المجلس الاعلى للامتحانات والتقويم التربوى الذى أصبح حسب المادة (٧) من القرار الجمهورى هو مجلس ادارة المركز الذى يتولى رئاسته وزير التعليم.

مهام المركز:

تتلخص مهام المركز القومى للامتحانات والتقويم التربوى كما حددتها المادة رقم (٣) من القرار الجمهورى فيما يلى:

أولا: وضع المعايير الخاصة لقياس وتقويم مختلف مستويات المعرفة والمهارات والجوانب الوجدانية للطلاب واعداد انظمة الامتحانات بما فى ذلك اسلوب تقدير الدرجات.

ثانيا: اعداد نظم الامتحانات كما يأتى:

أ- للشهادات العامة فى جميع المراحل التعليمية للتعليم قبل الجامعى بما يتلائم

مع أهداف المناهج الدراسية وطرق التدريس والوسائل التعليمية المتبعة.

ب- المستوى الرفيع للكشف عن قدرات الطلاب على التعليم الجامعى أو العالى ونظم اختبارات قدراتهم الخاصة.

ثالثا: متابعة المستوى الكيفى للامتحانات على اختلاف انواعها ومستوياتها والتحقق من سلامتها وكفائتها فى تقويم الطلاب.

رابعا: اجراء البحوث لتطوير وتحسين ورفع مستوى التقويم وادواته للطلاب فى جميع مستويات التعليم.

خامسا: التدريب على وضع الامتحانات وتصحيحها وإدارتها للموجهين والمعلمين والباحثين وغيرهم واعداد اختبارات مستوى الكفاية الفنية والمهنية لهم لتحقيق هذا الغرض.

سادسا: ابداء الرأى والمشورة وتقديم خبرة المركز فى نطاق اهدافه للجامعات والمعاهد العليا والفنية ومديريات التربية والتعليم بالمحافظات ولجميع الهيئات المهتمة بشئون التعليم بالداخل أو الخارج.

مهام المركز منذ عام ١٩٩٠*:

فى مجال مراقبة جودة التعليم:

١- تشكيل فرق لتقويم جوانب العملية التعليمية بالمدارس المختلفة فى كافة المراحل التعليمية على مستوى الجمهورية.

٢- اعداد تقارير عن أحوال هذه المدارس متضمنة الايجابيات والسلبيات وجهات الاحالة التى تنفذ ذلك.

٣- اعداد دليل للمقوم، وكذا قاعدة بيانات خاصة بمدارس المحافظات بجميع أنواعها على مستوى الجمهورية، والتى تم تخزينها فى الحاسب الآلى بالاضافة الى تسجيل كافة البيانات التى تتعلق بالمدارس السابق تقويمها لإمكان متابعتها.

٤- اعداد نموذج لتقويم المبانى المدرسية وصدور قرار الاستاذ الدكتور وزير التعليم رقم (٨) بتاريخ ١٩٩٥/٦/٢٢ باشتراك السادة المهندسين والفنيين المساعدين

* فؤاد أبو حطب ١٩٩٩ : (٦٦٠-٦٦٣)

بجهاز الخدمة الوطنية بالقوات المسلحة فى أعمال تقويم المبانى المدرسية بالتعاون مع المركز القومى للامتحانات والتقويم التربوى.

٥- تجريب المشروع الريادى للتقويم الشامل للمدرسة.

فى مجال تطوير الامتحانات:

١- اعداد مواصفات الاوراق الامتحانية من خلال لجان علمية متخصصة ابتداء من الصف الثالث الابتدائى وحتى شهادة الثانوية العامة.

٢- تشكيل لجان تقويم امتحانات الثانوية العامة خلال انعقادها.

٣- تقويم امتحانات الشهادة الاعدادية والصفين الثالث والخامس الابتدائيين وسنوات النقل على مستوى الجمهورية.

٤- اعداد ادلة تقويم الطالب فى جميع المواد الدراسية من الصف الثالث الابتدائى وحتى الصف الثالث الثانوى، وكذا أدلة المستوى الرفيع.

٥- اعداد بنوك اسئلة للمواد الدراسية المختلفة من الصف الثالث الابتدائى وحتى الصف الثالث الثانوى.

فى مجال التجريب والبحث:

١- تحليل نتائج امتحانات الثانوية العامة بمصر سنويا.

٢- تصميم تجربة لمحاكاة عمل المركز على مستوى الصف الرابع الابتدائى.

٣- تقويم منهج الصف الرابع الابتدائى والانشطة التعليمية والإدارة المدرسية واداء المعلم والمبنى المدرسى.

٤- تجريب بنوك الاسئلة على عينات ملائمة، وتحليل المفردات باستخدام الطرق التقليدية والطرق التى تعتمد على النماذج الرياضية الحديثة (مثل نموذج راش).

٥- اعداد قاعدة بيانات عن الاختبارات التربوية والنفسية فى مصر.

٦- اعداد دراسة قومية لتقويم برامج اعداد المعلم فى مصر.

٧- مشروع انشاء اختبار قومى للذكاء والقدرات العقلية.

٨- مشروع تطوير الاهداف التربوية والتعليمية.

فى مجال العمليات والمعلومات:

١- تصميم وبناء منظومة آلية باستخدام الكمبيوتر لادارة عمليات رصد الدرجات وتدقيقها ثم استخراج بيان بدرجات الطالب الناتج (ملصقة ترفق مع استمارته الدالة على النجاح).

٢- تصميم وبناء منظومة الية لبنوك الاسئلة من خلالها يتم تسجيل الاسئلة ومواصفاتها ومؤشراتها الاحصائية وبناء امتحان بشكل آلى طبقا لمواصفات دقيقة.

٣- تصميم وبناء منظومة آلية تقوم بعملية التحليل الاحصائى لعينات الشهادات العامة وخاصة الثانوية العامة.

٤- تصميم وبناء منظومة الية لتخزين واسترجاع معلومات التقويم والامتحانات.

٥- تجريب الميكنة الكاملة للامتحانات فى اعداد الامتحان وطباعته وتصحيحه واستخراج النتائج وتحليلها.

فى مجال التدريب والاعلام:

١- اعداد دراسة جدوى لتنظيم برامج تدريبية على مستوى الجمهورية لجميع العاملين بالتربية والتعليم بما يتفق مع تطوير التعليم مع الاهتمام خاصة بالتعليم الاساسى.

٢- عقد دورات تدريبية للمشاركين فى وضع أدلة تقويم الطالب ومفردات بنوك الاسئلة.

٣- عقد دورات تدريبية للمعلمين لتدريبهم على تقويم العمليات العقلية العليا للتلاميذ وتصميم الاسئلة التى تستخدم فى المواد الدراسية المختلفة لتحقيق هذا الهدف.

٤- الاستفادة بالخبرة الاجنبية فى مجال التقويم التربوى وقد انفتح المركز على مختلف الثقافات: تركيا، بريطانيا، فرنسا، الولايات المتحدة.

٥-عقد دورات للموجهين ورؤساء الاقسام التعليمية بالمرحلة الابتدائية للتدريب على الاساليب المتطورة فى الامتحانات فى مختلف المواد الدراسى والمجالات العملية

والانشطة واعداد مدربين للقيام بمهام التدريب محليا.

٦-المشاركة فى البرامج التدريبية لمنظمة اليونيسيف الخاصة بمدارس المجتمع.

٧- عقد لقاءات منتظمة مع قيادات التعليم فى المدريات التعليمية ومع المسئولين عن التوجيه والادارة المدرسية والمعلمين ومجالس الاباء والمعلمين للتعريف بتجارب المركز قبل البدء فيها واعداد المواد الإعلامية التى توضح اهمية استخدام اساليب التقويم المتطورة فى استثارة القدرات العقلية العليا للتلاميذ.

فى مجال النشر والنشاط العلمى:

يصدر المركز القومى للامتحانات دورية علمية متخصصة هى (المجلة المصرية للتقويم التربوى). كما يخطط الان (عام ١٩٩٦) لعقد المؤتمر الذى سيعقده الاتحاد الدولى للتقويم التربوى - والمركز عضو فيه - بالقاهرة عام ١٩٩٨ (١٢: ٦٦٣).

هذا ومازال المركز القومى للأمتجانات والتقويم التربوى يقوم بمهامه المنوطة به والمرتبطة بلائحته.

الفصل الثانى

مدخل إلى الإحصاء النفسى

أشرنا فى الفصل السابق إلى أن القياس النفسى يعتمد فى جوهره على البيانات الكمية التى تتوافر عن السلوك الإنسانى، وهى، كالبيانات الكيفية يجب أن تتاح للأخصائى النفسى كى يصدر أحكامه على هذا السلوك ويتخذ القرارات العملية حوله.

وبالطبع فإن البيانات الكمية التى تستند إليها عملية التقويم النفسى قد تكون أدق من البيانات الكيفية، ويتوقف ذلك على طبيعة الظاهرة النفسية موضع البحث والدراسة، ويجب على الأخصائى النفسى ألا يتعسف فيفرض على هذه الظاهرة طبيعة غير طبيعتها، ومن ذلك أننا حين نقارن بين شخصيتين فى «طول القامة» فإن من الأفضل أن نصف الشخص الأول مثلا بأن طوله ١٦٠سم ونصف الشخص الآخر بأن طوله ١٨٠ بدلا من تقول أن الشخص الثانى أطول من الأول، فالعبارة الأولى استخدمت لغة الكم، وهى أقرب إلى المقام وإلى طبيعة الظاهرة المقيسة، ولكننا عند المقارنة من شخصين فى القدرة على القيادة، يكون وصفنا اللفظى المفصل المعتمد على بصيرة الباحث أجدى وأفضل.

وتحتاج البيانات التى يحصل عليها الباحث إلى تحليل، ويستخدم علم الإحصاء فى تحليل البيانات الكمية خاصة أما البيانات الكيفية فلها طرقها الخاصة.

ونقدم فى هذا الفصل عرضاً موجزاً لطرق تحليل البيانات الكمية مصنفة تبعا لمستوى المقياس المستخدم، وقد سبق أن أشرنا إلى أن مستويات القياس هذه أربعة وهى المقاييس الاسمية، ومقاييس الرتبة ومقاييس المسافة ومقاييس النسبة وسوف نقتصر على الطرق المستخدمة فى الإحصاء الوصفى خاصة كما يتمثل فى الاتجاهات الرئيسية الثلاثة، النزعة المركزية والتشتت والعلاقة ، مع ملاحظة أن الطرق الإحصائية التى تصلح للمستويات الدنيا تصلح أيضاً للمستويات العليا والعكس غير صحيح*.

* يمكن لمزيد من التفصيل للطرق الإحصائية المتقدمة الرجوع إلى الكتاب الآتى: آمال صادق، فؤاد أبو حطب: مناهج البحث وطرق تحليل البيانات فى العلوم النفسية والتربوية والاجتماعية، الانجلو المصرية ١٩٩١ .

تحليل بيانات المقاييس الأسمية

أشرنا إلى أن المقاييس الاسمية تستخدم الأعداد لتشير إلى أفراد أو فئات ، وفى حالة استخدام الفئات، فإننا نقوم «بعد» عدد الحالات التى تقع فى كل فئة، ويطلق على هذا العدد «التكرار» وتوجد طرق إحصائية عديدة للتعامل مع التكرارات المصنفة إلى فئات، أما الأعداد التى تشير إلى الفئات ذاتها فلا تخضع لأي تحليل إحصائى لأنها فى الوقع محض رموز وقد يشير المقياس الاسمى إلى ما يسمى القيم المنفصلة وفى هذه الحالة ينطبق على أسئلة الاختبار إذا كانت الإجابة عليها من نوع «الكل أو لا شىء» مثل «إجابة صحيحة أو خاطئة» أو «نعم أو لا» وعوملت منفصلة، أي كل سؤال على حدة.

ويحتاج الباحث كما أشرنا فى تحليله لبيانات المقاييس إلى استخدام الطرق الإحصائية فى تحديد الاتجاهات الرئيسية الثلاثة: النزعة المركزية، والتشتت، والعلاقة. وقد لا تتضح هذه الاتجاهات جميعا فى المقاييس الاسمية، بينما نجدها كلها فى مقاييس الرتبة، ومقاييس المسافة والنسبة كما سوف تلاحظ.

النسب والنسب المئوية:

عادة ما يستخدم فى تحليل التكرارات النسب والنسب المئوية، والمثال التالى يوضح ذلك:

لاحظ أحد المدرسين اختلاف عدد التلاميذ والتلميذات فى فصلين من فصوله، كما يتبين من الجدول التالى:

جدول (١)

	تكرار التلاميذ	تكرار التلميذات	المجموع
الفصل (أ)	٣٠	٣٠	٦٠
الفصل (ب)	٣٠	٢٠	٥٠

إنه لا يستطيع أن يستنتج من الجدول السابق أن الوضع النسبى للذكور فى الفصلين متساو ما دام عدد التلاميذ الذكور متساو، ويفيد فى المقارنة أن يقوم بتحويل التكرارات المتضمنة فى الجدول السابق إلى نسب أو نسب مئوية كما يلى:

جدول (٢)

	نسب التلاميذ	نسب التلميذات	النسبة المئوية للتلاميذ	النسبة المئوية للتلميذات
الففصل (أ)	,٥	,٥	٥٠٪	٥٠٪
الفصل (ب)	,٦	,٤	٦٠٪	٤٠٪

وقد حصلنا على نسبة الذكور مثلا بقسمة تكرارهم على المجموع الكلى لتلاميذ الفصل (أى ٣٠ ÷ ٦٠) أما النسبة المئوية فقد حصلنا عليها بضرب النسبة السابقة × ١٠٠ ، ولعلك لاحظت بعد هذا التحويل أن الذكور يمثلون نسبة أو (نسبة مئوية) أكبر من الإناث فى الفصل (ب) ، بينما تتساوى النسبتان فى الفصل (أ) وتدل بالطبع النسبة ٠,٥ أو النسبة المئوية ٥٠٪ على نقطة التوسط أو النزعة المركزية ، وقد يستخدم فى قياس النزعة المنوال mode ويدل على أكثر القيم شيوعا فى التوزيع التكرارى ، وعلى الرغم من أنه مقياس مباشر للنزعة المركزية إلا أنه أقل دقة من المتوسط والوسيط اللذين سنشير إليهما فيما بعد.

المدى المطلق :

المدى المطلق هو أحد مقاييس التشتت وأكثر بساطة وسهولة ، وهو فى جوهره مقياس للسمة ، ويحسب مباشرة بتحديد الفرق بين أكبر عدد وأقل عدد فى البيانات المتوافرة ، وهو بهذا المعنى يصلح للاستخدام مع المقاييس الاسمية ، فإذا كانت أرقام جلوس التلاميذ تمتد من ١ إلى ٥٠٠ فإن المدى المطلق فى هذه الحالة يصبح ٤٩٩ ، ويمكن استخدامه بنفس الطريقة مع درجات الاختبارات.

وقى يستخدم فى قياس تشتت بيانات المقاييس الاسمية النسبة والنسب المئوية أيضاً. فكلما ازدادت هذه النسبة تطرفا بالزيادة أوالنقص دل ذلك على درجة البعد عن نقطة التوسط (٥٠٪)

معامل الارتباط بين بيانات المقاييس الاسمية :

يدل معامل الارتباط على مدى الاتفاق بين متغيرين ، أى ما إذا كانا يختلفان معها ، وفى أى اتجاه ، وإلى أى حد؟ أو بمعنى آخر هو النزعة إلى اقتران التغير فى ظاهرة بالتغير فى ظاهرة أخرى.

ويمتد معامل الارتباط بين (+١) والذى يعنى وجود ارتباط موجب كامل بين المتغيرين، (فالصفر) الذى يعنى عدم وجود ارتباط بينهما واستقلالهما عن بعضهما تماما، (-١) الذى يعنى وجود ارتباط سالب كامل بين المتغيرين، والواقع أن الحصول على معاملات الارتباط الكاملة، سواء بالموجب أوالسالب، نادر الحدوث إن لم يكن مستحيلا، ولذلك فإن أغلب معاملات الارتباط بين المتغيرات فى المجالات النفسية والتربوية والاجتماعية عبارة عن كسر عشرى (أكبر من الصفر وأقل من الواحد الصحيح) يدل على وجود علاقة جزئية، وقد يكون هذا الكسر العشرى موجبا أو سالبا.

وتوجد طرق إحصائية كثيرة لقياس العلاقة بين التكرارات المصنفة إلى فئات المقاييس الاسمية، وهى تنتمى إلى ما يسمى الإحصاء الابارامترى، ولايتسع المقام لعرض هذه الطرق ويمكن الرجوع فى ذلك إلى مرجع متخصص (١)، ونكتفي بعرض طريقة واحدة هى ما يسمي معامل ارتباط فاى.

مثال: حصل أحد المعلمين على البيانات التالية من عدة فصول يقوم بالتدريس لها تتألف من الجنسين (الذكور والإناث) صنفوا جميعا فى ضوء نجاحهم أو فشلهم الدراسى كما يلى:

جدول (٣)

	التلاميذ	التلميذات	المجموع
ناجحون	(٣٥)	(٣٧)	٧٢
فاشلون	(١٤)	(٣٤)	٤٨
المجموع	٤٩	٧١	١٢٠

وأراد هذا المعلم أن يحلل البيانات الاسمية فى الجدول السابق ليستخلص منه ما إذا كانت توجد علاقة بين الجنس والنجاح المدرسى، إنه عليه فى هذه الحالة أن يحسب معامل ارتباط فاى، وذلك بتحويل التكرار الثنائى للأعداد بين القوسين فى الجدول السابق إلى اللغة الرمزية كما يلى: **جدول (٤)**

التلاميذ	التلميذات
أ	ب
جـ	د

ثم تطبق المعادلة الآتية:

$$\text{ر ف} = \frac{\text{أ د} - \text{ب جـ}}{\sqrt{(\text{أ} + \text{ب})(\text{جـ} + \text{د})(\text{أ} + \text{جـ})(\text{ب} + \text{د})}}$$

وبالتعويض عن رمز المعادلة فإن:

$$\text{ر} = \frac{(٣٥ \times ٣٤) - (٣٧ \times ١٤)}{\sqrt{(٧٢ \times ٤٨ \times ٥٩ \times ٧١)}} = +٠,١٩$$

وهو معامل ارتباط موجب ولكنه منخفض، وقد لا تكون له دلالة إحصائية(*) أى لا يدل على وجود علاقة بين الجنسين والنجاح المدرسى.

تحليل بيانات مقاييس الرتبة

يمكن الحصول على اتجاه النزعة المركزية للبيانات الرتبية باستخدام الوسيط، وعلى اتجاه التشتت باستخدام نصف المدى الربيعى، وعلى معامل الارتباط كمقياس للعلاقة باستخدام معامل ارتباط للرتب.

الوسيط:

يمكن تعريف الوسيط بأنه الرتبة التى تقسم التكرارات إلى نصفين متساوين، وفيما يلى مثال يوضح طريقة حسابه:

(*) مفهوم الدلالة الإحصائية ينتمى إلى الاحصاء الاستدلالى، ولا يتسع المقام فى هذا الكتاب لشرحه ويمكن للقارئ الرجوع إلى بعض المصادر المتخصصة فى الإحصاء النفسى والتربوى والاجتماعى- راجع المرجع رقم (١) .

جدول (٥)

فئات الدرجات	الحدود الحقيقية للفئات	التكرار	التكرار المتجمع الصاعد
٥ – ٩	٤,٥ – ٩,٥	٤	٤
١٠ – ١٤	٩,٥ – ١٤,٥	١	٥
١٥ – ١٩	١٤,٥ – ١٩,٥	١	٦
٢٠ – ٢٤	١٩,٥ – ٢٤,٥	١٠	١٦ فئات الوسيط
٢٥ – ٢٩	٢٤,٥ – ٢٩,٥	٣	١٩
٣٠ – ٣٤	٢٩,٥ – ٣٤,٥	٥	٢٤
٣٥ – ٣٩	٣٤,٥ – ٣٩,٥	٣	٢٧
٤٠ – ٤٤	٣٩,٥ – ٤٤,٥	٠	٢٧
٤٥ – ٤٩	٤٤,٥ – ٤٩,٥	١	٢٨
	المجموع	٢٨	

ولحساب الوسيط من بيانات الجدول السابق يتبع الباحث الخطوات الآتية:

١ – تحديد رتبة الوسيط أى النقطة التى ينقسم عندها التكرار إلى نصفين بقسمة المجموع الكلى للتكرارات على ٢ أى أن رتبة الوسيط = $\frac{٢٨}{٢}$ = ١٤

٢ – دراسة التكرارات المتجمعة الصاعدة (العمود الأخير فى الجدول) لتحديد الفئة التى يقع فيها الوسيط. وهى فى هذه الحالة الفئة (٢٠ – ٢٤).

٣ – تطبيق المعادلة التالية:

$$\text{الوسيط} = \text{الحد الأدنى الحقيقى لفئة الوسيط} + \left[\frac{\text{رتبة الوسيط} - \text{التكرار السابق لفئة الوسيط}}{\text{تكرار فئة الوسيط}}\right] \times \text{مدى الفئة}$$

على النحو الآتى:

$$\text{الوسيط} = ١٩,٥ + \frac{١٤ - ٦}{١٠} \times ٥ = ٢٣,٥$$

نصف المدى الربيعى:

لعلك لاحظت أن الوسيط هو نقطة واحدة تقسم التكرار الكلى أى مجموع الأفراد أوالحالات إلى نصفين، ويمكن لهذه الفكرة أن تمتد منطقيا إلى تقسيم التكرار إلى أى عدد من الأقسامو فمثلا يمكن أن ينقسم إلى أربعة أقسام باستخدام ثلاث نقط التى تسمى فى هذه الحالة الربيعيات، أو إلى عشرة أقسام باستخدام تسع نقط تسمى العشيريات، أو مائة قسم باستخدام ٩٩ نقطة تسمى المئينيات، وتلتقى هنا بفكرة الربيعيات التى يعتمد عليها حساب نسف المدى الربيعى كمقياس إحصائى لتشتت البيانات الرتبية.

ولايختلف حساب الربيع عن حساب الوسيط، إلا فى نقطة البداية وعلى هذا حساب الربيع الأدنى أو الأول من بيانات الجدول السابق كما يلى:

١ – تحديد رتبة الربيع الأول بقسمة المجموع الكلى للتكرار على (٤) وضربها فى (١) أى أن:

$$\text{رتبة الربيع الأول أو الأدنى} = \frac{٢٨ \times ١}{٤} = ٧$$

وهم ربع الأفراد الذين يفترض فيهم أن تقع درجاتهم أدنى من نقطة الربيع الأول أو الأدنى، أما الأرباع الثلاثة الأخرى فتقع درجاتهم أعلى منها.

٢ – دراسة التكرارات المتجمعة الصاعدة لتحديد الفئة التى تقع فيها نقطة الربيع الأول وهى فى هذه الحالة الفئة (٢٠ – ٢٤)

٣ – تطبيق المعادلة التالية التى لا تختلف فى صيغتها عن معادلة حساب الوسيط.

$$\text{الربيع الأول} = \text{الحد الأدنى الحقيقى لفئة الربيع الأول} + \left[\frac{\text{رتبة الربيع الأول – التكرار المتجمع السابق لفئته}}{\text{تكرار فئته}}\right] \times \text{مدى الفئة}$$

أى أن:

$$\text{الربيع الأول} = ١٩,٥ + \frac{٢ \times ن}{١٠} \times ٥ = ٢٠$$

أما الربيع الثانى فتتحدد من رتبته على النحو التالى بافتراض ن =

المجموع الكلى للتكرار.

رتبة الربيع الثانية أو الأوسط = $\frac{ن \times ٢}{٤} = \frac{ن}{٢}$ = الوسيط

وهم نصف الأفراد الذين يفترض فيهم أن تقع درجاتهم أدنى من نقطة الربيع الثانى أوالوسيط، وبالطبع فإن النصف الثانى تقع درجاتهم أعلى منها.

ومعنى ذلك: أن الربيع الثانى هو الوسيط وسبق لنا حسابه.

أما الربيع الثالث أو الأعلى فيتحدد من بيانات الجدول السابق كما يلى:

١ – رتبة الربيع الثالث أو الأعلى = $\frac{ن \times ٣}{٤} = \frac{٢٨ \times ٣}{٤}$ = ٢١

وهم ثلاثة أرباع الأفراد الذين يفترض فيهم أن تقع درجاتهم أدنى من نقطة الربيع الثالث أو الأعلى، أما الربع الباقى فتقع درجاتهم أعلى منها.

٢ – الفئة التى يقع فيها الربيع الثالث هى ٣٠ – ٣٤

٣ – الربيع الثالث = ٢٩,٥ + ($\frac{٢١ - ١٩}{٥}$) × ٥ = ٢١,٥

ولحساب نصف المدى الربيعى كمقياس إحصائى لتشتت البيانات الرتبية تعتمد على نصف الفرق بين الربيعيين الأول والثالث وهو ما يسمى نصف المدى الربيعى.

نصف المدى الربيعى = $\frac{\text{الربيع الثالث – الربيع الأول}}{٢}$

= $\frac{٢١,٥ - ١٩}{٢}$ = ٥,٧٥

معامل ارتباط الرتب:

يستخدم فى القياس الإحصائى للعلاقة بين متغيرين من طبيعة رتبيةمعامل ارتباط الرتب الذى ابتكره عالم النفس الإنجليزى تشارلز سبيرمان، ويوضح الجدول التالى طريقة حسابه.

مثال :

حصل أحد الباحثين على ترتيب ١٥ تلميذا فى مادتى التاريخ والجغرافيا، حيث تدل الرتبة (١) على التلميذ الأول والرتبة (١٥) على التلميذ الأخير فى كل مادة على حدة، مع ملاحظة أن التلاميذ الذين تساوت رتبتهم حصلوا على متوسط الرتب المتتالية، أى أن التلاميذ د، س حصلا فى التاريخ على رتبة ١٢,٥ وهى متوسط الرتبتين ١٢، ١٣ والتلاميذ أ. ز، ن، حصلوا جميعا على الرتبة (٨) وهى متوسط الرتب ٧، ٨، ٩.

جدول (٦)

التلاميذ	ترتيب التاريخ	ترتيب الجغرافيا	فروق الرتب (ف)		مربع فروق الرتب (ف٢)
أ	١١	٨	+	٣	٩,٠٠
ب	٤	٦	−	٢	٤,٠٠
جـ	٩	٥	+	٤	١٦,٠٠
د	١٠	١٤	−	٤	١٦,٠٠
هـ	١٤,٥	١٥	−	٠,٥	٠,٢٥
و	١٤,٥	١٢	+	٢,٥	٦,٢
ز	١٢,٥	٨	+	٤,٥	٢٠,٢٥
ح	١	٣	−	٢	٤,٠٠
ط	٣	١	+	٢	٤,٠٠
ى	٧	٤	+	٣	٩,٠٠
ك	٦	١٠	−	٤	١٦,٠٠
ل	٢	٢		صفر	٠٠,٠٠
م	٥	١٣	−	٨	٦٤,٠٠
ن	٨	٨		صفر	٠٠,٠٠
س	١٢,٥	١١	+	١,٥	٢,٢٥

مجـ ف = صفر مجـ ف٢ = ١٧١,٠٠

ويحسب معامل ارتباط الرتب بالمعادلة الآتية:

$$ر_ت = ١ - \left[\frac{٦ \times مجـ ف^{٢}}{ن (ن^{٢} - ١)}\right]$$

$$= ١ - \frac{٦ \times ١٧١}{١٥ \times ٢٣٤} = ٠,٦٩٦$$

وهو معامل ارتباط موجب مرتفع نسبياً، ويدل على وجود علاقة بين التحصيل فى كل من التاريخ والجغرافيا أى أن التلميذ الذى جاء ترتيبه مرتفعا فى التاريخ يحتمل أن يكون كذلك فى الجغرافيا والعكس صحيح.

تحليل بيانات مقاييس المسافة والنسبة

يحصل الباحث على اتجاه النزعة المركزية للبيانات التى توفرها مقاييس المسافة والنسبة باستخدام المتوسط الحسابى ، وعلى اتجاه التشتت بحساب الانحراف المعيارى، وعلى معامل الارتباط باستخدام طريقة بيرسون التى تسمى طريقة الارتباط التتابعى، (أو طريقة حاصل ضرب العزوم).

المتوسط الحسابى:

المتوسط الحسابى كمقياس للنزعة المركزية لبيانات مقاييس المسافة والنسبة، يعتمد على تقسيم الدرجات أو الكميات - وليس التكرارات كما هو الحال فى المتوسط - إلى نصفين متساويين، ويوضح المثال التالى ذلك:

مثال: طبق أحد المعلمين اختبارا فى فهم القراءة على تلاميذ أحد فصول الصف الرابع الابتدائى فحصل على البيانات الآتية:

جدول(٧)

التلاميذ	أ	ب	جـ	د	هـ	و	ز	ح	ط
الدرجات	١	٣	٤	٥	٧	٥	٣	١	٧

وللحصول على المتوسط الحسابى كمقياس إحصائى للنزعة المركزية يقوم المعلم بجمع درجات التلاميذ، وهو فى هذه الحالة ٣٦، وجمع عدد التلاميذ وهو فى هذه

الحالة ٩، ويحسب المتوسط الحسابى بالمعادلة الآتية:

المتوسط الحسابى = ـــــــــــ (أو) م = ـــــــــ = ـــــ = ٤

مجموع الدرجات / عدد الأفراد — مجـ س / ن — ٣٦ / ٩

الانحراف المعيارى:

يعتمد حساب الانحراف المعيارى كمقياس إحصائى لتشتت بيانات مقاييس المسافة والنسبة على حساب المتوسط الحسابى أولا، ثم حساب الفروق أو الانحراف عن هذا المتوسط، ونظرا لأن مجموع هذه الفروق يساوى صفرا، يستحيل علينا أن نحسب متوسط هذه الفروق، ولذلك يلجأ الإحصائيون إلى تربيع كل فرق أو (انحراف) للتخلص من العلامات الجبرية السالبة، ثم نحسب بعد ذلك متوسط مربعات الانحراف بقسمتها على عدد الأفراد، ثم نستخرج الجذر التربيعى لمتوسط مربعات الانحراف، ويكون هذا الجذر التربيعى هو الانحراف المعيارى.

ويمكن حساب الانحراف المعيارى للمثال السابق، وذلك بطرح المتوسط الحسابى (= ٤)، من كل درجة من الدرجات السابقة فتحصل على الانحراف عن المتوسط (ح)، ثم نربعها (ح٢) كما يلى:

جدول (٨)

التلاميذ	أ	ب	جـ	د	هـ	و	ز	ح	ط
ح	-٣	-١	صفر	+١	+٣	+١	+١	-٣	+٣
ح٢	٩	١	صفر	١	٩	١	١	٩	٩

مجموع مربعات الانحرافات = ٤٠

$$\text{الانحراف المعيارى} = \sqrt{\frac{\text{مجموع مربعات الانحرافات عن المتوسط}}{\text{عدد الأفراد}}}$$

$$\text{أو ع} = \sqrt{\frac{\text{مج ح٢}}{\text{ن}}} = \sqrt{\frac{٤٠}{٩}} = ٢,١١$$

معامل الارتباط التتابعى لبيرسون:

يستخدم فى حساب العلاقة بين مقاييس المسافة والنسبة الطريقة العامة فى حساب معامل الارتباط والتى ابتكرها عالم الإحصاء الإنجليزى كارل بيرسون ، ويوضح المثال التالى طريقة الحصول على هذا المعامل.

مثال: طبق أحد المعلمين اختبارين يقيس أحدهما تحصيل تلاميذ الصف الرابع من التعليم الأساسى فى مادة الحساب (س) ويقيس الآخر القدرة العددية (ص) لدى هؤلاء التلاميذ، ويوضح الجدول الآتى البيانات التى حصل عليها وخطوات حساب معامل الارتباط.

جدول (٩)

التلاميذ	درجات اختبار التحصيل الحسابى (س)	درجات اختبار القدرة العددية (ص)	س٢	ص٢	س ص
أ	٢	٥	٤	٢٥	١٠
ب	٣	٧	٩	٤٩	٢١
جـ	٥	٦	٢٥	٣٦	٣٠
د	٧	١٠	٤٩	١٠٠	٧٠
هـ	٨	١٢	٦٤	١٤٤	٩٦
ن = ٥	مج س = ٢٥	مج ص = ٤٠	مجـ س٢ = ١٥١	مجـ ص٢ = ٣٥٤	مجـ س ص = ٢٢٧

ويمكن حساب معامل الارتباط مباشرة من بيانات الجدول السابق بتطبيق المعادلة الآتية:

$$ر = \frac{(ن\ مجـ\ س\ ص) - (مجـ\ س \times مجـ\ ص)}{\sqrt{[ن\ مجـ\ س^{٢}) - (مجـ\ س)^{٢}] \times [ن\ مجس\ ص^{٢}) - مجـ\ ص)^{٢}]}}$$

وبالتعويض عن رموز المعادلة السابقة نجد أن:

$$ر = \frac{(٥ \times ٢٢٧) - (٢٥ \times ٤٠)}{\sqrt{[(٥ \times ١٥١) - ٦٢٥][(٥ \times ٣٥٤) - ١٦٠٠]}} = ٠,٩١$$

وهو معامل ارتباط مرتفع جداً، ويد على أن التلميذ الذى يحصل على درجات عالية فى اختبار التحصيل الحسابى يحتمل احتمالا كبيراً أن يكون على درجة عالية فى اختبار القدرة العددية، والعكس صحيح.

مراجع الباب الأول

١ -آمال صادق، فؤاد أبو حطب: مناهج البحث وطرق التحليل الاحصائى فى العلوم النفسية والتربوية والاجتماعية - الانجلو المصرية - ١٩٩١ .

٢- أحمد زكى صالح: الأسس النفسية للتعليم الثانوى: القاهرة : دار النهضة العربية، ١٩٧٢.

٣ - أحمد زكى صالح: علم النفس التربوى: القاهرة: النهضة المصرية، ١٩٧٢.

٤- أمين محمد صبرى نور الدين: فاعلية استخدام الإختبار الموائم باستخدام الحاسب فى تقدير قدرة الأفراد وتحديد الخصائص السيكومترية للمقياس- رسالة دكتوراه غير منشورة - كلية التربية - جامعة عين شمس ٢٠٠٢ .

٥ - جابر عبدالحميد: التقويم التربوى والقياس النفسى، دار النهضة العربية، ١٩٨٣ .

٦ - سعيد يامشموش، السيد محمد خيرى، يحيى مهنى: التقويم التربوى الرياض: منشورات دار الفيصل الثقافية، ١٩٨٠ .

٧ - صلاح الدين محمود علام: تحليل البيانات فى البحوث النفسية والتربوية . دار الفكر العربى، ١٩٨٥ .

٨ - رايتسون، ج. جاستمان، ج. و. روبنز، (ترجمة : محمد محمد عاشور وآخرين): التقويم فى التربية الحديثة . القاهرة: الأنجلو المصرية ، ١٩٦٥ .

٩ - رمزية الغريب: القياس اللابرمترى، مكتبة الأنجلو المصرية، ١٩٨٥ .

١٠ - رمزية الغريب: التقويم والقياس النفسى والتربوى، القاهرة: الأنجلو ١٩٧٠ .

١١ - عبدالله السيد عبدالجواد: المؤشرات التربوية واستخدام الرياضيات فى العلوم الإنسانية، أسيوط، ١٩٨٣ .

١٢ - فؤاد أبو حطب، القدرات العقلية . القاهرة، الأنجلو المصرية ط٤، ط ٥، ١٩٩٦ .

١٣- فؤاد أبو حطب: الإرتقاء بمستوى خريج التعليم الجامعى والعالى فى اطار مفهوم الجودة الكلية لمواجهة تحديات المستقبل، المجلس القومى للتعليم ١٩٩٩ . فى: المجلة المصرية للدراسات النفسية، العدد التذكارى ٣١ : ٢٠٠١

١٤ - فؤاد أبو حطب، آمال صادق، علم النفس التربوى، الأنجلو المصرية (ط ٣)، ١٩٨٤، (ط ٤) ١٩٩٦، (ط ٨) ٢٠٠٤ الأنجلو المصرية.

١٥ – فؤاد البهى السيد: علم النفس الإحصائى وقياس العقل البشرى، دار الفكر العربى، ١٩٧٨ .

١٦ أ – محمد خليفة التونسى : التقويم والتقييم، مجلة العربى، يونيه ١٩٨٦ .

١٦ ب – محمد خليفة التونسى: أضواء على لغتنا السمحة، كتاب العربى، ١٩٨٦ .

١٦ جـ – محمد خليفة التونسى: الاشتقاق من المشتق، مجلة العربى، ١٩٨٦ .

١٧ – محمد عبدالسلام أحمد: القياس النفسى والتربوى، القاهرة، النهضة المصرية، ١٩٦٠ .

١٨ – نعيم عطية: التقييم التربوى الهادف، بيروت : منشورات دار الكتاب اللبنانى، ١٩٧٠ .

١٩ – يوسف محمود الشيخ، جابر عبدالحميد جابر: سيكولوجية الفروق الفردية، القاهرة: النهضة المصرية، ١٩٦٤ .

20. Abouserie, R & Selwyn, N. (1997). The Computerized Versus Standard Short Term Memory Tests: The Possible Effects Of Computer Anxiety And Experience, School Of Education. Cardiff University. Wales.

21. Alkhader, O., Clark, D., & Anderson, N. (1998) Equivalence And Predictive Validity Of Paper - And- Pencil And Computerized Adaptive Formats Of The Differential Aptitude Tests. Journal Of Occupational And Organizational Psychology, 713, 205-218.

22. Allport, G. W. Pattern And Growth In Personality, New York: Holt, Rinehart, 1961, (1937).

23. Anastasi, A. Differential Psychology (4rd Ed.) New York: Macmillan, 1953.

24. Anastasi, A. Psychological Testing New York, Macmillan Publishing, 5th Ed.) 1982.

25. Bloom, B., Et Al, Handbook On Formative And Sammative Evaluation Of Student Learning N.y. McGraw Hill, 1991 .

26. Bloom, B. Taxonomy Of Educational Objectives, London: Longmans, 1953.

27. Bordon, E. S. Psychological Counseling. New York: Appleton, 1955.

28. Cook, W. W. The Function Of Measurement In The Facilitation Of Learning. In Lindquist, E. F. (Ed.) Educational Measurement Amer. Council On Educ. Washington, D.. C., 1951, 3-64.

29. Combs, C.h. Theory And Method Of Social Measurement. In: Festinger, L. And Katz (Eds.) Research Methods Of Behavioral Sciences. New Delhi Ed., 1970.

30. Cronbach, J. L. Essentials Of Psychological Testing, New York: Harper (4th Ed.), 1981.

31. Fliege, H., Becker, J.; Walter, O.; Bjorner, J.; Klapp, B.; & Rose, M. Development Of A Computer- Adaptive Test For Depression (D-cat). Quality Of Life Research, Vol. 14, 2277-2291.(2005).

32. Forbey, J. D. & Ben -porath, Y. S. Computerized Adaptive Personality Testing: A Review And Illustraton With The Mmpi-2 Computerized Adaptive Version. Psychological Assessment, Vol. 19, 12-24. (2007).

33. Gronlund, Measument And Evaluation In Teaching, New York. Macmillan, (4th Ed) 1984.

34. Hambleton, R., Zaal, J. & Pieters, Jo P. M. Computerized Adaptive Testing: Theory, Applications, And Standards. In R. K. Hambleton, & J. N. Zaal (Eds.), Advances In Educational And Psychological Testing And Applications.(Pp 341-366) Boston : Kluwer Academic (1994) .

34.a. Jensen, A. R. Bias In Mental Testing, London, 1980 .

35. Krathwohl, D. R., Et Al. Taxonomy Of Educational Objectives. Ii Affective Domeuin, 1964.

36. Mcbride, J., & Martine, J. Reliability And Validity Of Adaptive Ability Tests In A Military Setting. In D. J. Weiss (Ed.), New Horizons In Testing. (Pp. 223-236) New York: Academic Press.(1983) .

37. Meijer, R. & Ering, M. Computerized Adaptive Testing: Overview And Introduction. Applied Psychological Measurement, 23. 187-194. (1999).

38. Moreno, K. Sgall, D. & Hetter, R. The Use Of Computerized

Adaptive Testing In The Military. In R. F. Dillon. (Ed.) Handbook Of Testing. (Pp. 204-219) London: Greenwood Press.(1997).

39. Murphy, K., & Davidshofer, C. O. Psycological Testing: Principles And Applications. 3rd., Edit New Jersey: Prentice- Hall. (1994) .
40. Nunnally, J. C. (Jr.)introduction To Psychological Measurement, New York: Mcgraw-hill, 1970.
41. Rogers, C. R. Counseling And Psychotherapy. Cambridge, Riverside Press, 1951.
42. Rogers, C.r. Client-centered Therapy, Cambridge, Mass.: Riverside Press, 1951.
43. Stevens, S.s. (Ed.) Handbook Of Experimental Psychology. New York: Wiley, 1951.
44. Sundberg, N. Assesment Of Persons. N. J. Prentice- Hall Inc, 1977.
45. Thorndike, R. And Hague, E. Measurement And Evaluation In Psychology And Education, New York: Wily, 1955.
46. Thorndike, R. And Hague, E. Measurement And Evaluation In Psychology And Education, New York: Wily 1955 .
47. Tyler, R. W. The Function Of Measurement In Improving Instruction. In: Lindquist, E. F. (Ed.) Educational Measurement, 1951, 47 - 67.
48. Tyler, L. C. The Work Of The Counselor (2nd Ed.) New York : Appleton-century, 1961.
49. Urry, V.tailored Testing A Successful Application Of Latent Trait Theory. Journal Of Educational Measurement, 14, 181-196. (1977).
50. Wang, T., & Kolen, M. Evaluating Comparability In Computerized Adaptive Testing: Issues, Criteria And An Example. Journal Of Educational Measurement, 38, 19-49.(2001).

الباب الثانى
شروط الاختبارات النفسية

يشمل هذا الباب الفصول التالية:

- فن اعطاء الاختبارات
- ثبات الاختبارات
- صدق الاختبارات
- تصحيح الاختبارات ،تفسير الدرجات

الفصل الثالث

فن إعطاء الاختبارات

ماذا يقصد بتقنين الاختبار؟

أشرنا فى الفصل الأول إلى تعريف للقياس النفسى يشير إلى ضرورة التعبير الصريح عن قواعده، وهذا ما يتضمنه مفهوم «التقنين Standardization . فحين نقول أن هذا المقياس «مقنن» فإن ذلك يعنى فى جوهره أنه لو استخدمه أفراد مختلفون يحصلون على نتائج متماثلة، ويتطلب هذا بالطبع توحيد إجراءات تطبيق الاختبار وتصحيحه، فإذا كانت الدرجات التى يحصل عليها الأفراد المختلفون قابلة للمقارنة فإن شروط الاختبار وظروف إعطائه يجب أن تكون موحدة للجميع، وهذا الشرط ما هو إلا تطبيق خاص للحاجة إلى الشروط والظروف المضبوطة فى جميع الملاحظات العلمية، ومثل هذا التقنين يجب أن يشمل المواد المستخدمة وحدود الزمن والتعليمات الشفوية والتحريرية التى تعطى للمفحوصين والأمثلة التوضيحية، وطرق تناول أسئلة المفحوصين واستفساراتهم ، وكل ظرف أو حالة لاترتبط بما يتناوله الاختبار ويمكن أن تؤثر فى أداء المفحوصين فى الاختبار كلما كان ذلك ممكنا.

ويوفر تقنين قواعد القياس للباحثين فى علم النفس ظرفا ملائما لسهولة «الاتصال»، لأن من أهم خصائص العلم وجود نوع من الاتصال الجيد بين العلماء، بحيث يستطيع الباحث أن يقارن بين نتائجه ونتائج زملائه الذين يتناولون نفس المشكلات بالدراسة، فإذا افترضنا أن أحد الباحثين أجرى تجربة على أثار التوتر فى استجابات القلق وسجل نتائجه فى عبارة مثل أن «التوتر يؤدى إلى ظهور القلق عند الأطفال» فإن مثل هذه النتيجة لاتفيد فى «الاتصال» الجيد لأن الباحثين الآخرين قد يختلفون حول ما يقصده «بظهور القلق»، بل قد يصعب عليهم، إن لم يستحل ، تكرار نفس التجربة للتحقق من صحة الفرض، ويمكن تحقيق نوع من الاتصال الأكفأ لو استطاع الباحث أن يلاحظ القلق باختبار «مقنن» أى محددة قواعد إعطائه وشروط تطبيقه.

وتتأثر عملية إعطاء الاختبار بعوامل متعددة، منها ما يتصل بمن يعطى الاختبار وما يفعله، ومنها ما يتصل بمن يقوم بأداء معين على هذا الاختبار وما يحيط به من

ظروف ومؤثرات، ومدها ما يتصل بالاختبار نفسه وما يتميز به من خصائص وما يقصد إلى أن يقيس، وإذا كانت هناك عوامل وشروط ينبغى أن تتوفر فى الاختبار نفسه حتى يكون أداة قياس يعتمد عليها ويطمأن إليها. وهى ما سنتناوله فى الفصول التالية، فإن هناك شروطا وظروفا تؤلف ما نسميه فن إعطاء أوتطبيق الاختبارات على الفاحص أن يراعيها ويحققها، وهى التى من شأنها أن توفر، أول ما توفر، الجو النفسى ـ الاجتماعى المناسب للمفحوص بحيث يأتى أداؤه فى الاختبار ، كما أراد واضعه، وكما يريد الفاحص، دقيقا وصادقا، ليكون أساسا متينا لقرار تزيد احتمالات دقته وجدواه.

ولعل مسألة تقنين الاختبارات وغيرها من أدوات التقويم النفسى هى التى جعلت علماء النفس على درجة كبيرة من الحساسية لمسألة أخلاقيات المهنة، وقد ظهر هذا الاهتمام لأول مرة فى أوائل الخمسينيات حين أصدرت الجمعية الأمريكية لعلم النفس مجموعة من المعايير الأخلاقية التى يجب على الباحثين السيكولوجيين مراعاتها (١٥) ثم سرعان ما اتبعت ذلك بإصدار كتيب التوصيات الفنية الخاصة بالاختبارات النفسية والأساليب التشخيصية (١٦)، وقد تطورت هذه الوثائق وتعدلت طوال السنوات الثلاثين الماضية وصدرت من كل منها طبعات جديدة (١٧، ١٨، ١٩، ٢٠)، ونحن فى حاجة إلى ترشيد هذه الخبرة والممارسة السيكولوچية فى مصر، وقد عرض هذه الضرورة محمد خليفة بركات (١١) فى المؤتمر الثانى لعلم النفس فى مصر عام ١٩٨٦، وتقوم الجمعية المصرية للدراسات النفسية بهذه المسئولية بالاشتراك مع رابطة الاخصائيين النفسيين (رانم) وعقد لذلك العديد من الجلسات.

واجبات الفاحص ومسئولياته فى إعطاء الاختبار:

١ - **التدريب على إعطاء الاختبار:** من الاختبارات ما يسهل اعطاؤه فلايحتاج إلى تدريب خاص للفاحص، ومنها ما يحتاج إلى تدريب ومران قد يمتد بضعة شهور، والاختبارات الفردية بصفة عامة تحتاج إلى تدريب أكثر من الاختبارات الجماعية، وإن كانت هناك بعض الاختبارات الجماعية التى تتطلب عناية كبيرة فى إعطائها، وكذلك فإن الاختبارات الإسقاطية تحتاج إلى تدريب خاص قد يصل فى بعض الجامعات إلى الحصول على دبلومة متخصصة فيها، وهو ما

لاتحتاجه الاستخبارات أو الاستبيانات، وتحتاج المقابلة أو الملاحظة تدريبا متخصصاً لا تحتاجه بعض وسائل التقرير الذاتى، وإذا كانت كل مسئولية الفاحص تنحصر فى قراءة مجموعة من التوجيهات والتعليمات المطبوعة، فإن أى شخص ذا ضمير يقظ حساس، وليس فى شخصيته ما يحدث تهديدا للآخرين، يمكن أن يقوم بها بنجاح، أما إذا كانت الضرورة تقتضى أن يسأل المفحوص فرديا، وأن تستعمل أسئلة تتبعية إذا كان السؤال الأول غير واضح، فإن هذا يقتضى توفر قدر كبير من الخبرة والمهارة، أى تدريباً خاصاً وطويلا بما يتفق مع درجة صعوبة الموقف الاختبارى، ودقة القرار الذى يتخذ فى ضوء بيانات الاختبار.

وتشير بعض المؤسسات المسئولة عن طبع الاختبارات وتوزيعها، وعلى الخصوص NFER فى إنجلترا ETS(*) فى أمريكا أن تتوافر فى من يشترى الاختبارات شروط ومؤهلات معينة ـ قد تصل إلى الدكتوراه فى علم النفس قبل السماح ببيعه أو توزيعه، ناهيك عن القيود الصارمة على التوزيع والنشر التى تفرضها هذه المؤسسات وغيرها على الاختبارات حتى لا تفقد وظائفها الأساسية فى التقويم النفسى. ولك أن تقارن هذا بما يشيع فى بلادنا من فوضى حيث تباع الاختبارات وتشترى، بل وتصور، دون رقيب، بل يصل الأمر إلى حد إعادة طبعها كاملة فى ملاحق رسائل الماجستير والدكتوراه وفى هذا كله خروج على الميثاق الأخلاقى السيكولوجى لاستخدام الاختبارات.

٢ – **ألفة الفاحص بالاختبار:** من أهم الواجبات التى ينبغى أن يراعيها الفاحص أن يتيح الفرصة لكل مفحوص ليظهر قدرته، وليحصل على نتائج يمكن مقارنتها بنتائج مفحوصين آخرين، فعندما تقوم عملية انتقاء ما على نتائج اختبار من الإختبارات فإنها لن تكون عملية عادلة ما لم يتم اختبار الأشخاص جميعا بطريقة موحدة ومتشابهة تماما.

ولهذا كان من الضرورى أن يعرف الفاحص الذى سيستعمله، بل ويألفه إلى أبعد درجة ممكنة، إذا كنا نهدف إلى تحقيق عملية اختبار دقيقة منصفة. ومما يساعده

(*) هما: (NFER) National Foundation for Educational Research
(ETS) Educational Testing services.
على التوالى.

على هذه المعرفة والألفة أن يقرأ كراسة التعليمات التى تصحب الاختبار عادة بعناية وأن يلتزم بما فيها التزاما تاما.

٣ - **الالتزام بالاتجاه العلمى:** على الفاحص أن يلتزم اتجاها علميا محايدا أثناء عملية الاختبار، ذلك لأن المختبر غالبا ما يكون مهتما بالمفحوصين، وقد يميل إلى أن يراهم يحسنون صنعا فى الاختبار الذى يعطيهم إياه، فينزع إلى أن يساعدهم بنية حسنة، بطريقة مباشر أو بطرق خفية، وهذا أكثر حدوثا فى مواقف الاختبار الفردية، ولكن على المختبر أن يلاحظ هذا فى نفسه، وأن يتجنبه، ولاشك أن مما يساعد على هذا أن يلتزم المختبر بالتعليمات المصاحبة للاختبار، كما أشرنا، فإذا ذكرت التعليمات مثلا أن سؤالا ما يكرر مرة واحدة، فلا يكرره المختبر أكثر من مرة واحدة فى أى ظرف من الظروف.

وكذلك على المختبر أن يلاحظ أى تعبيرات للوجه أو كلمات التشجيع أو الإشارات التى يراد بها حث همة المفحوص، هذا كله يضعف من قيمة الاختبار ومن تقنينه، وعلى هذا فمن المهم بالنسبة للفاحص أن يجعل هيئته وسمته وتعبيرات وجهه وإشاراته غير كاشفة، على ألا يعطى هذا المفحوص إحساسا بأن المختبر لايكترث به، بل لابد أن يحرص المختبر على أن يوحى للمفحوص ويؤكد له فى صمت اهتمامه بكل ما يقول.

٤ - **المحافظة على العلاقة بين الفاحص والمفحوص:** يلفتنا هذا إلى أهمية المحافظة على العلاقة بين الفاحص والمفحوص فى موقف الاختبار. فإذا كان مطلوبا من المفحوص أن يظهر أفضل ما عنده فى الاختبار، أى أنه ينبغى أن يظهر باستمرار رغبته فى التعاون مع المختبر فإنه من الضرورى أن تقوم علاقة إنسانية تتسم بالاطمئنان والثقة والقبول المتبادل بين طرفى هذه العملية - الاجتماعية فى أساسها - أى علاقة المفحوص والمختبر.

ونسارع فنشير إلى أنه ليست هناك قواعد محددة ولا نصائح تتبع لإقامة ذلك النوع من العلاقة بين الفاحص والمفحوص، ولكننا نستطيع أن نقول بصفة عامة أن الشخص الذى يألف الناس ويألفونه، والذى يوحى بالثقة وبالطمأنينة فى نفوس الآخرين، يستطيع أن يجد من الوسائل ما يساعده على إنشاء تلك العلاقة بسماتها التى حددناها.

وهناك علامات تدل على سوء العلاقة الاختبارية، أو على أن الفاحص غير موفق فى إقامة العلاقة المطلوبة بمن يختبرهم، ونذكر بعضا من هذه العلامات حتى يستطيع الفاحص الفطن أن يتعرف عليها ويدركها ثم يستجيب لها بما يؤدى إلى اختفائها أو الإقلال منها، من هذه العلامات عدم انتباه المفحوصين أثناء إلقاء التعليمات، أو الانتهاء قبل أن يحين الموعد المحدد، أو عدم الاستقرار، أو التماس أخطاء فى الاختبار، أو إظهار اليأس من إمكان الأداء أوالعمل فى الاختبار.

٥ - **مسئولية الفاحص الأولى اختيار الاختبار الملائم للمفحوص:** من الممارسات السيئة والخطيرة أن يعطى الفاحص للمفحوص اختبارا للقدرة مثلا لا يناسبه ثم يؤدى فيه أداءا سيئا لأسباب غير مرتبطة بالأداء الاختبارى ذاته، فاختبار المتاهة مثلا لا يجب أن يعطى لمفحوصين يعانون من نقص فى التحكم فى حركات اليد، وبالمثل لايجب أن يعطى اختبار للميول المهنية يلائم طلاب الجامعات لطلاب مرحلة التعليم الأساسى ما لم يكن قد أجرى تقنين جديد له يؤكد صلاحيته للمستوى التعليمى الجديد، ويشمل ذلك بالطبع التأكد من أن تلاميذ مرحلة التعليم الأساسى يفهمون تعليماته وأسئلته.

ومن المعلوم فى علم القياس النفسى وفنونه التطبيقية أن حصول المفحوص على درجة مضللة فى اختبار غير ملائم أسوأ من عدم الحصول على هذه الدرجة على الإطلاق، ومسئولية اختيار الاختبار الملائم للمفحوص تظل فى النهاية مسئولية الفاحص، وتبقى مسئوليته هذه حتى ولو طلب منه تطبيق اختبار بذاته من سلطة خارجية فعليه أن يستخدم مهارته وكفاءته فى الحكم على ملاءمة الاختبار،و ويصدق هذا على مساعدى الباحثين الذين يقومون بجمع البيانات، ولعلنا فى هذا الصدد لانبالغ إذا قلنا أن بحثا أو دراسة أو فحصا يعتمد على بيانات زائفة أو سيئة أو مضللة أسوأ للعلم وأضل للممارسة العملية من عدم إجراء بحوث على الإطلاق.

طريقة إعطاء الاختبار:

١ - **ظروف عملية الاختبار:** هناك ظروف مادية لابد أن تتوفر فى الموقف الاختبارى لما لها من آثار عميقة على سلوك الأشخاص من الفاحصين وبالتالى على نتائج اختباراتهم، من هذه الظروف التهوية والإضاءة ولهما أثر واضح على أداء

الأشخاص فى الاختبارات، سواء كان هذا الأداء كتابة أو عملا. وكذلك من الضرورى توفير أماكن جلوس تساعد على القيام بهذا الأداء، من حيث الاتساع وإمكان وضع الأدوات على مناضد بحيث يسهل تناولها واستعمالها وهكذا.

أما ترتيب أماكن الجلوس فلابد أن يراعى فيه أن يساعد المفحوصين على الاستماع الجيد والسليم للإرشادات والتوجيهات، وهنا نشير إلى أن الصالات الكبيرة الواسعة لاتكون مناسبة للاختبارات الجماعية إلا إذا توفر العدد الكافى من المشرفين على أن يوزعوا توزيعا مناسبا، وحتى مع توفر العدد الكافى من المشرفين فلا تزال هناك صعوبة لابد أن نتغلب عليها وهى أن المفحوصين فى صالة ضخمة قد يحجمون عن السؤال أو الاستفسار من المشرفين خوفا ورهبا، وعلى هذا، ينبغى عند إجراء اختبار جماعى فى صالة كبيرة أن يشجع المفحوصون على أن يسألوا ويستفسروا عما يعن لهم أو يلتبس عليهم فيما يتصل بما بين أيديهم من اختبار.

أما إذا انتقلنا إلى ظروف المفحوص نفسه، من الناحية الجسمية والنفسية، فيجب أن نعطيها قدرا كبيرا من الاهتمام، وأول ما يلحظ هو اختيار الوقت المناسب له من الناحية الجسمية والنفسية، لأن إعطاء الاختبار فى ظروف جسمية أو انفعالية غيرمناسبة، لاتجعل أداءه عينة صادقة وممثلة لسلوكه.

ومما يساعد على تحقيق هذه الظروف توفير قسط مناسب من الراحة قبيل إعطاء الاختبارات، ثم أثناءها بصورة منظمة دون ارتباك أو فوضى، وكذلك مراعاة أن لا تعطى الاختبارات عقب فترة من المجهود الجسمى أو العقلى الشاق أو عقب التعرض لحالة انفعالية حادة، إلا إذا كان الغرض من القياس ذلك.

أما إذا اضطررنا، بسبب ما، أن نعطى اختبارات فى ظروف غير مناسبة فيجب أن نعتبر نتائجها أولية وأن ننظر إليها نظرة ناقدة وأن نأخذها بغير قليل من التحفظ والحذر..

٢ - **إدارة الجماعة وضبطها:** لا يعطى الاختبار إلا لأفراد نتوقع أن يفهموا توجيهات وتعليمات معينة، كما نتوقع أن يسلكوا وفق هذه التوجيهات والتعليمات التى يوجهها الفاحص اليهم، ولذا كان أساس الاختبار الجماعى، من ناحية الفاحص، هو ضبط الجماعة وإدارتها بطريقة تمكنه شخصيا من أداء واجباته المتصلة بإعطاء

الاختبار بدقة وحكمة، ثم تمكن الجماعة من إظهار أحسن أداء ممكن على هذا الاختبار.

ومما يحقق للفاحص النجاح فى هذا أن يقف أمام الجماعة بحيث يرى كل أعضائها وبحيث يستطيع كل أعضائها أن يروه ويسمعوه، ويلقى تعليماته بعد أن يستقر ويسكن الجميع فى أماكنهم، ويلاحظ أن تكون التعليمات بسيطة واضحة وأن يلقيها الواحدة أثر الأخرى بهدوء وأناة بحيث يكون كل جزء وكل فقرة من التعليمات متميزة ليتجنب سوء الفهم والخلط من جانب المفحوصين، وإذا كان انتباه الجماعة له كاملا، وهو ما ينبغى أن يحرص عليه، فإنه لن يحتاج إلى تكرار شىء مما يقول.

ومن الممكن أن يلجأ الفاحص إلى التمثيل ليزيد التعليمات وضوحًا، ما لم يكن هناك نص فى دليل الاختبار أو تعليمات موجهة إلى الفاحص نفسه تنص على منع التمثيل، ففى هذه الحالة عليه أن يتجنب التمثيل والتوضيح. ومن أمثلة التوضيح مثلا أنه إذا كانت التعليمات تشير إلى قلب الصفحة الأولى من الاختبار، فلابأس من أن يمسك المختبر الاختبار عاليا ثم يقلب الصفحة، أو أن يشير إلى المكان الذى ينبغى أن تكتب فيه البيانات الخاصة بالمفحوص، وهكذا.

ومن المهم أن يكون واضحاً أن للمفحوصين الحق دائما فى أن يستفسروا ويسألوا، ولعله من الأفضل أن يتوقع المختبر الأسئلة ويحاول الإجابة عنها قبل أن تسأل.

هل يتحقق ضبط الجماعة المفحوصة ويتيسر إجراء الاختبار فيها عن طريق اتباع أسلوب صارم جامد من الفاحص؟ لاشك أن مثل هذا الأسلوب يجعل أمر ضبط الجماعة وإدارتها سهلا لامشقة فيه ولاعناء على الفاحص إلا أن مثل هذا الأسلوب يخلع على جو الاختبار طابعاغير إنسانى، ومن المهم أن لا ننسى فى الموقف الاختبارى فى أى موقف اجتماعى بل وإنسانى بصفة عامة، أنه إذا كان ضبط الجماعة ضرورة لتسيير أمورها وتحقيق مطالبها، فإن من الضرورى أيضًا أن يحس أعضاء الجماعة بأن من يوجههم أو يشرف عليهم إنسان يهتم بهم ويحرص عليهم وليس شخصاً رسميا يقيم بينهم وبين نفسه أستاراً تجعلهم ينفرون منه مما يمتد أثره إلى ما بين أيديهم من عمل يطلب منهم أداؤه، والموقف الاختبارى على أى حال، لا يتطلب إلا قدرا من الضبط يحققه قدر يسير من الرسمية فى العلاقات بين المفحوصين

والفاحص. وقد تزيد درجة الرسمية هذه أو تنقص حسب أغراض الاختبار وظروفه، ففى بعض الاختبارات التحصيلية - وخاصة فى مواقف الامتحانات - قد يحتاج الأمر لدرجة كبيرة من الرسمية وإلا حدثت الفوضى وشاع الغش، ولعل ما يحدث فى وقتنا الحاضر فى الامتحانات المدرسية من مشكلات نظام يؤيد هذا القول، وهو على أية حال أمر يحتاج لدراسة مستقلة.

ومن المهم أن نشير إلى أن الفاحص لابد أن يرجع إلى فطنته وتقديره ويعتمد على فصاحته فى كثير من الأحيان أثناء الموقف الاختبارى، فإذا نشأت ظروف اضطرارية أثناء الاختبار، كأن يمرض مفحوص فجأة، مثلا، فهنا لابد أن يحتكم الفاحص إلى تقديره الخاص، مراعيا أن الهدف من أى اختبار هو الحصول على معلومات مفيدة ودقيقة عن الشخص وله، وليس مجرد إجراء اختبار كأنه عمل روتينى والفراغ منه على أى حال وبأية صورة.

٣ - **التعليمات للمفحوص**: إن إعطاء التعليمات والإرشادات أهم مسئولية يتحملها الفاحص، ذلك لأنه من الضرورى فى الاختبارات المقننة أن يعطى المختبر التعليمات كما وردت فى الاختبار دون زيادة أو نقص، فعلى الرغم من أن الاختبارات تقيس الفروق الفردية فإن طرقها ووسائلها تكون مرسومة بحيث تلغى التفاوت والاختلاف والتباين فى كل خاصية من خواص المفحوص فيما عدا تلك الخاصية التى يريد أن يقيسها الاختبار بالفعل، وعلى هذا فإنه من الأهمية بمكان بالنسبة للفاحص أن تتساوى جميع الظروف - قدر الإمكان - بين المفحوصين فيما عدا ذلك العامل أو تلك الخاصية التى يريد الاختبار قياسها.

وعادة ما تسمح التعليمات نفسها بالأسئلة والاستفسارات، ولكن عندما يجيب الفاحص على الأسئلة عليه ألا يتعرض لتوضيح شىء أكثر مما فى التعليمات، إذ عليه أن يتذكر دائما أن نتائج مفحوصية قد تقارن بنتائج مفحوصين آخرين فى أماكن وأوقات أخرى.

وعلى أى حال فإن الاختبار الجيد يأخذ فى الاعتبار كما أشرنا أكبر قدر من الأسئلة المتوقعة ويرتب إجابات وحلول لها حتى تكون طريقة إعطائه مقننة، وبدون ذلك فإن هذه الأسئلة قد يفهمها البعض على وجه بينما يفهمها البعض الآخر على

وجه آخر، ومن الضرورى دائما تحديد المطلوب من المفحوصين بصورة واضحة محددة إلى أبعد حد مستطاع.

وفى بعض الأحيان تكون طريقة اتباع المفحوص للتعليمات المقننة مقصودا بها أن تؤثر فى درجته، فإذا أعطيت هذه التعليمات بطرق مختلفة ترتب على ذلك اختلافا فى الأداء لم يكن فى الحسبان.

ومن المسائل الهامة التى تتصل بتقنين تعليمات الاختبار ما يتصل بالتخمين، ويظهر التخمين بصورة أقوى وأوضح فى اختبارات القدرة واختبارات التحصيل حيث يسأل المفحوص: «إذا لم أعرف الإجابة فهل من الممكن أن ألجأ إلى التخمين؟» وفى بعض الأحيان تتضمن توجيهات الاختبارات نصوصًا تبين الإجابة على مثل هذا السؤال، ولكن حتى مع وجود مثل هذه النصوص فإن الأمر لا يخلو من ابهام يتطلب الوضوح، وأنه مما يتعارض مع أصول إعطاء الاختبارات أن يتطوع الفاحص بإعطاء نصح وتوصية من عنده لم تنص عليه توجيهات الاختبار.

ما هو الموقف السليم بالنسبة للتخمين فى الاختبارات؟ يمكن أن نلخص ما ينبغى أن نعمله فى مواجهة التخمين فيما يلى:

(١) على الفاحص أن يؤكد ما تشير إليه تعليمات الاختبار، فإذا كانت هذه التعليمات تشجع التخمين أكده وشجع عليه أما إذا كانت تنص على عدم الاعتماد على التخمين فعليه أن يظهر هذا.

(ب) أما إذا لم تكن فى دليل الاختبار إشارة إلى التخمين فعلى الفاحص أن ينصح المفحوصين بأن يستعملوا ذكاءهم وتقديرهم فيما يختص بما إذا كان من الممكن أن يخمنوا، عندما تواجههم أسئلة أو عناصر من الاختبار لايكونون واثقين من صحة الإجابة عنها، بحيث تكون تخميناتهم مستنيرة، فلا يختارون إلا ما يعتقدون أنه أقرب إلى الصواب، ولاتكون المسألة كلها فى يد الصدفة والتخمين غير المعقول.

(جـ) من الضرورى أن يتعلم أولئك الذين سوف تطبق عليهم الاختبارات وخاصة من الطلاب، أن الإسراف فى التخمين فى اختبارات القدرة واختبارات التحصيل ليس فى صالحهم على الإطلاق.

(د) من المعروف أنه قد وضعت معادلات لتصحيح أثر التخمين فى الاختبارات، وعلى الفاحص أن يلجأ إلى استخدام المعادلة المناسبة عند تصحيح الاختبارات ليفسر الدرجة التى يحصل عليها المفحوص تفسيرا معقولا.

استثارة الدوافع لأخذ الاختبار:

هناك حقيقة مهمة على كل مشتغل بالقياس النفسى أن يعيها وأن يعمل وفقا لها أولا وهى أن الشخص الذى نختبره لابد أن يهتم بنتائج الاختبار الذى يعطى له وأن تكون لهذه النتائج قيمة عنده، ففى القياس النفسى - الذى يختلف عن القياس المادى - نجد أن المفحوص يضع نفسه فى المقياس، وما لم يهتم شخصيا بهذه العملية فإننا لا نستطيع قياسه وإن لم يتحقق هذا فإننا لايمكن اختباره بدرجة نطمئن فيها على ما نحصل عليه من بيانات عنه، وإذا كانت الدوافع لاتمثل جانبا له أهمية في أنواع القياس المادى المختلفة، مثل الطول والوزن، فإنه لها أهميتها الكبرى فى القياس النفسى، وحتى يكون القياس النفسى دقيقا لابد أن يقبل المفحوص على الاختبار ويتحرك نحوه، ولهذه كانت استثارة دوافع المفحوص للاختبار مسألة جديرة باهتمام الفاحصين وعنايتهم.

(١) الحوافز التى ترفع الدرجة: الحوافز التى تساعد على الأداء فى الاختبارات النفسية متعددة ومتنوعة، فمنها مثلا الرغبة فيها، ومنها ذلك الحماس الذى ينتج من اندماج الذات والاستغراق الذى يعبر عن الرغبة فى الاحتفاظ باحترام الذات واحترام الآخرين، ومن هذه الحوافز أيضًا مجرد حب العمل والاشتغال بحل عناصر اختبار ما، فمثل هذا الحب والاهتمام يستثير الشخص إلى بذل مجهود أكبر فيه، ولا نستطيع أن نغفل حافزا يقوم بدور مهم فى بعض الأحيان فى دفع المفحوصين وتحريكهم للأداء على اختبار ما، ذلك هو الرغبة فى إطاعة السلطة وتنفيذ التزامها كما تتمثل فى الفاحص وما يوجهه إلى المفحوصين من تعليمات وتوجيهات، وهذه الرغبة كما هو واضح متصلة بحاجات اجتماعية عميقة منها مثلا الحاجة إلى الاعتراف والتقدير، وكذلك الحاجة إلى الضبط والنظام.

وتظهر الحاجة إلى استثارة حوافز المفحوص ورفع دوافعه إلى الدرجة الملائمة فى اختبارات القدرة واختبارات التحصيل بصفة خاصة، وذلك لأن أهم ما يعنينا من

هذين النوعين من الاختبارات هو الحصول على أحسن أداء ممكن عليها من المفحوصين، أو الحصول على أعلى إنتاجية ممكنة فيهم إذا استعملنا تعبيرا ذا مسحة اقتصادية.

والذى لاشك فيه أن مراعاة ما ذكرناه، وما سنناقش فيما يلى من شروط وظروف فى الموقف الاختبارى، مما يساعد على تحقيق ذلك المستوى الأنسب من الدوافع لأخذ الاختبار والاستمرار فيه بجهد متوازن مناسب. وقد بذلت محاولات متعددة لاستعمال أنواع بسيطة من المكافآت لرفع مستوى الأداء على اختبارات القدرة بنوع خاص، ومن أمثلة هذه المكافآت البسيطة الجوائز، والنقود، وأحاديث التشجيع، وقد أشارت الأبحاث التى أجريت فى هذا الصدد (٢١، ٢٥، ٢٨) أن مثل هذه المكافآت لم تحدث أي تحسن فى درجات اختبارات القدرة عن الدرجات التى حصل عليها المفحوصون الذين أخذوا الاختبارات نفسها فى الظروف العادية لإعطاء الاختبارات، ومما يلاحظ أنه من الممكن أن تتحسن الدرجات على الاختبار عندما تحول دوافع الشخص من التركيز على مثل تلك المكافآت البسيطة إلى الاهتمام بدرجات الاختبار التى يحصل عليها.

وحتى فى الاختبارات التى تعطى لأغراض البحث العلمى وليس لاتخاذ قرار بشأن المفحوصين فإنه لابد أن يعطى الفاحص أو القائم بالبحث للمفحوصين أسباباً لأخذ الاختبار لها قيمة شخصية لهم، وبدون هذا فإن أداءهم أو مشاركتهم فى الاختبار تكون عملا لا معنى له، مما ينعكس على نوع الأداء وبالتالى على البيانات التى نقيم عليها البحث كله..

(٢) **الدوافع التى تخفض الدرجة**: أما الدوافع التى تؤدى إلى خفض الدرجة على الاختبار فنستطيع أن نحدد منها ما هو متعمد مقصود، وما هو متصل بالرغبة الزائدة فى الأداء الممتاز والتفوق وما يصاحب هذا من قلق وتوتر.

وليس عجيبا أن يعمد مفحوص إلى أن يكون أداؤه سيئا بحيث تكون درجته على الاختبار أقل كثيراً مما يستطيع ، ولكنه يفعل هذا لأنه، مثلا، لايريد أن يحرم من عضوية جماعة يحبها ويحرص على أن يبقى فيها وحصوله على درجة عالية فى الاختبار كفيل بأن يبعده عن جماعته إلى جماعة أخرى أو لأنه يتوقع أن تضاف إليه

أعباء وواجبات إضافية نتيجة حصوله على درجة عالية فى الاختبار وهكذا.

أما الدافع الشديد إلى الإمتياز والتفوق فى الأداء على اختبار ما فقد يؤدى إلى التأثير السيىء على هذا الأداء، كما أن التوتر الزائد يؤدى إلى أخطاء فى الأداء كما نلاحظ فى الاختبارات اللغوية أو الاختبارات التى تتطلب دقة حركية وتآزراً حسياً حركيا بنوع خاص.

ومن المهم أن نشير إلى أن قلق الاختبار، كما أظهرت الأبحاث التى أجريت عنه، متعلم من معاملة الآباء والمدرسين للتلاميذ ، كما أنه متعلم من رفقاء التلاميذ وأصحابهم ويبدو أن قلق الاختبار موجود بدرجات تكاد تتساوى بين القادرين من التلاميذ وبين غير القادرين منهم كما أثبتت أبحاث ساراسون Sarason وزملائه (٣٢)، مستعملين فى قياس قلق الاختبار اختبارا وضعوه لهذا الغرض خاصة.

وقد وجدت بعض التجارب التى أجريت على قلق الاختبار علاقة سالبة بين القلق والأداء فى اختبارات التحصيل واختبارات الذكاء، مما يؤيد الفرض الذى يذهب إلى الأطفال القلقين فى موقف الاختبار يهبط مستوى أدائهم، ولكن، كما تشير انستازى (١٢ : ٣٦ – ٣٨) لاينبغى أن نأخذ هذه العلاقة السالبة بين القلق والأداء فى الإختبار على أنها حتما إشارة إلى علاقة سببية بينهما، ذلك لأنه من المحتمل أن ينشأ عند المفحوص قلق لأن أداءه فى الاختبار كان ضعيفا، وليس العكس. ومما تشير إليه نتائج الدراسات فى هذه الناحية، أن درجة بسيطة من القلق تساعد فى الأداء فى الاختبارات وفى التعلم، بينما المستوى العالى من القلق يكون له أثر معطل، وتشير النتائج أيضًا إلى أن التعليمات الموجهة إلى المفحوصين التى تتضمن توضيحاً لأهمية الاختبار ونتائجه للمفحوصين وتستدعى اندماجه فى الأداء فيه، هذا النوع من التعليمات يحدث أثرا إيجابيا على أداء المفحوصين الذين يكون لديهم مستوى منخفض من القلق عندما يبدأون الأداء فى الاختبار، بينما تحدث أثرا سيئًا عند أولئك المفحوصين الذين يبدأون بمستوى عال من القلق.

على أن العلاقة بين القلق وبين الأداء فى الاختبارات، وكذلك العلاقة بين القلق وبين التعلم، تعتبر من المسائل التى تحتاج إلى بحوث وتجارب عديدة حتى تتحدد بصورة أدق وأوضح، ويمكن الرجوع إلى بعض نتائج البحوث فى هذا الصدد إلى

بعض المراجع المتخصصة (٩).

ومن الطبيعى أن نتوقع أن كل موقف اختبارى يتضمن نوعا من التهديد أو من الاعتراف بقبول وجود سمة أو خاصية شخصية قد يكشف عنها الاختبار، مثل العصابية أو الميل إلى التسلط، لهذا، من المهم أن يؤكد الفاحص للمفحوص أن النتائج سوف تستخدم لصالحه، مع تأكيد الجوانب الإيجابية بصفة خاصة من هذا الاستخدام.

التهيئة لأخذ الاختبار:

(١) **إدراك المفحوص لمغزى الاختبار:** من المهم أن يرغب المفحوص فى أن تكون درجته فى الاختبار صادقة صحيحة، ذلك لأن هذه الرغبة ضمان لأن تكون الدرجة صادقة فعلا، ويساعد التعاون بين الفاحص والمفحوص على إيجاد هذه الرغبة، ولهذا كان أمراً يستحق أن يحرص المختبر على توفيره، وهو ليس عسير التحقق على أى حال، على أنه من المهم أن يكون واضحاً أن التعاون الذى نقصده بين المختبر والمفحوص ليس معناه أن تحال المسئولية كلها وتنقل إلى المفحوص، وإنما نقصد أن يجعل الفاحص المفحوص يحس باستمرار أنه مشارك له فى العمل مشاركة العضو العامل فى أى جماعة ولذا كان من الضرورى أن نتذكر أن الموقف الاختبارى المثالى هو الموقف الذى يكون المفحوص فيه شريكا فى اختبار نفسه، كما سبقت الإشارة عند الكلام عن دوافع المفحوص.

وعلى الفاحص أن يعطى المفحوص كل ثقته فيما يتصل بهدف الاختبار وأن يصور له الاختبار على أنه فرصة متاحة وأداة علمية ميسرة يعرف عن طريقها شيئا عن نفسه، تجعله أقدر على اتخاذ القرار المناسب له، وعلى فهم ما يتخذ من قرارات متصلة به سواء فى العلاج أو الإرشاد النفسى أو عند التقدم لكلية أو معهد أو وظيفة ما. وإذا كان الطبيب البارع يشرح لمريضه نوع الدواء الذى يعينه له والنتائج التى يتوقعها من تناوله الدواء. فإن المختبر الممتاز يجب أيضاً أن يجعل الهدف من أخذ الاختبار والنتائج التى تترتب عليه واضحة بالنسبة للمفحوص، ولاشك أن هذا الوضوح يجعل المفحوص أكثر حرصاً على أن يجعل الاختبار يعطى صورة دقيقة بقدر الإمكان عما يراد قياسه.

(٢) **دور المعلومات التمهيدية:** يزداد الاعتراف فى الوقت الحاضر بأهمية

المعلومات السابقة التى يعطيها الفاحص للمفحوصين فى تهيئتهم وإعدادهم للاختبار، ففى الإرشاد النفسى مثلا (٢٢) لم يعد إعطاء الاختبارات مسألة روتينية يقوم بها المفحوصين قبل أن تبدأ عملية الإرشاد النفسى، بل أصبح من الضرورى أن تسبق إعطاء الاختبارات مقابلات وغالبا ما يفضل مقابلتان أو ثلاثة، تتحدد فيها المشكلة ويناقش المختبر المفحوصين فى أنسب الاختبارات للمشكلة والغرض الذى من أجله يسعى هؤلاء المفحوصين إلى الإرشاد النفسى. هذا، وأن لمثل هذه المقابلات أثراً لايمكن إغفاله وهو أنها تساعد على الإقلال من قلق الاختبار الذى تحدثنا عنه.

(٣) دور التدريب على الاختبار: ظهرت فى السنوات الأخيرة برامج ومؤسسات لتدريب المفحوصين على أخذ الاختبارات النفسية، كما ظهرت فى الأسواق سلاسل من «الاختبارات المصغرة» يتدرب عليها المفحوصون قبل أن يتعرضوا لبرامج الانتقاء المهنى أو التعليمى. وازدادت هذه الظاهرة شيوعاً فى المجتمعات التى تلعب فيها الاختبارات دوراً حاسماً فى هذا المجال، وخاصة الولايات المتحدة الأمريكية.

وقد شغلت مسألة التدريب على الاختبارات النفسية الباحثين فى ميدان القياس النفسى منذ نشأته المبكرة، وأجريت تجارب مهمة للتحقق من فعالية التدريب على الاختبارات، ويمكن أن نلخص نتائج هذه البحوث فيما يلى (٢٣: ٧٨ - ٨٠).

(أ) تختلف برامج التدريب على الاختبارات فى طبيعتها، وتعتمد فعاليتها على التهيؤ المبدئى لدى المفحوص وعلى دوافعه، وقد لوحظ أن مجرد توافر الفرصة للمفحوص لأن يعطى الاختبار مرتين له بعض الأثر، وخاصة إذا لم تكن لديه فى الأصل فكرة عن الاختبار، كما لوحظ أيضاً أن مجرد تخصيص عدد من الساعات لإعطاء إشارات وتلميحات عن طريقة أخذ الاختبار وللتدريب عليه يمكن أن تزيد درجة المفحوص ما بين ٨، ١٠ نقاط.

(ب) تزداد فعالية التدريب كلما تحول من مجرد مراجعة سريعة إلى مقرر منظم، فعند تدريب بعض الطلاب على حل بعض المسائل الرياضية لمدة ٢١ ساعة من الدروس داخل الفصل و٢١ ساعة أخرى من الواجبات المنزلية أدى ذلك إلى زيادة الدرجة بمتوسط ١٦ نقطة فى أحد اختبارات الاستعداد الرياضى وحدث ما يشبه ذلك فى اختبارات الاستعداد اللغوى، ويزداد متوسط عدد نقاط الدرجة بزيادة ساعات

التقويم النفسى ـــــــ

فى التوصيات التى توجه إلى الفاحصين، وإنما الأمر يتضمن شخصين، فى حالة الاختبار الفردى مثلا، يمثل أحدهما، وهو المفحوص، شخصا عنده تصورات عن نفسه وعن الاخرين، واقعية أو غير واقعية، وتصورات عن الفاحص نفسه وعن الموقف الاختبارى، وعنده كذلك مخاوف وشكوك وتوقعات وآمال، وأنه لحامل ذلك كله إلى الموقف الاختبارى حيث يؤثر فيه من غير شك، أما الشخص الآخر فهو الفاحص، وهو إنسان عنده أيضًا توقعات وآمال ومخاوف وشكوك وإن اختلفت نوعا ودرجة عما عند المفحوص، بل وعنده تحيزات وتفضيلات لابد أن تنعكس وتؤثر فى كل مرحلة من مراحل الاختبار: إعطائه وتصحيحه وتفسيره.

ومن الطبيعى أنه عندما يأتى شخصان بهذه الحالة التى صورناها موجزين إلى موقف اختبار أو قياس أو تقويم فلابد أن نتوقع نشأة علاقات اجتماعية ونفسية متشابكة تجعل مهمة المختبر أكثر دقة ورهافة.

وفيما يلى نذكر أمثلة لهذا النوع من العلاقات تكفى لأن تلفت النظر إلى إمكان تنوع هذه العلاقات وإلى غموضها، وتساعد الفاحص فى دور التدريب أو الممارسة الفعلية على أن ينمى عنده حساسية لهذه العلاقات فى المواقف الاختبارية، بحيث يكتشفها بسرعة فيواجهها بصورة تجنبه خطأ أوتجاوزا يمكن تجنبه بالالتفات إلى مثل ما سنشير إليه، وتطبيقه فى مران طويل من مواقف اختبارية تجريبية وحقيقية.

قد يحس شخص محول ليجرى عليه اختبار معين أو عدة اختبارات أن السلطة التى حولته تستبد به وتميل إلى أن تسيطر عليه، ويكون رد الفعل عنده هو التمرد على هذه السلطة، وعلى الفاحص الذى يمثل «سلطة» أخرى بالنسبة إليه، ويظهر هذا التمرد المنقول من السلطة الأصلية إلى سلطة الفاحص فى مقاومات ومظاهر معينه أثناء الموقف الاختبارى، كأن يعطى مثلا معلومات يعتقد أنها مضللة عن نفسه، أو يعطى معلومات خاطئة عمدا، أو يظهر إهمالا وعدم اكتراث فى الإجابة أو الأداء أيا كانت صورته، أو يبدى اعتراضات كثيرة، وما إلى هذا فى مظاهر المقاومة والفشل فى العلاقة بين المختبر والمفحوص التى تكلمنا عنها من قبل، ولابد أن يدرك الفاحص هذا، خاصة فى المواقف الفردية، ويعمل على الإقلال من هذه المقاومة ومعالجتها بالاعتراف بوجودها وتقبلها وإبرازها للمفحوص، وتأكيد الضرر الذى سيلحق به من

أثر هذه المقاومة، ومظاهرها المعطلة على النتائج التى يعطيها الاختبار والتى هى فى النهاية لصالح المفحوص، وبعد ذلك يمكن أن يعطى الاختبار.

بل إن الشخص الذى يأتى بنفسه طالبا إجراء اختبار أو عدة اختبارات عليه، ربما لا يخلو هو أيضًا من صعوبة على الرغم من توفر الدافع لأخذ الاختبار عنده، ولو ظاهريا على الأقل، والأغلب أن تتوقع تعاونا كبيراً بين مثل هذا الشخص وبين الفاحص، ولكن قد تتعارض دوافع هذا المفحوص مع أهداف الاختبار، فإذا كان الدافع عنده هو أن يتعرف على بعض جوانب من شخصيته، فإنه فى الوقت نفسه يخشى أن يكشف الاختبار، أو الفاحص عن نواحى قصور عنده يحرص هو بكل وسيلة أن يبقيها خارج نطاق وعيه، وأكثر من هذا، فإن مجرد سعى المرء إلى أخذ اختبار، أو إلى المساعدة النفسية، فيه مساس باستقلاله الشخصى وإحساسه بقيمته وقدرته بحيث يزيد هذا عنده القلق إزاء الاختبار والفاحص جميعا.

وقد يكون الدافع الظاهر، أو المعلن، للسعى لأخذ اختبار غير الدافع الحقيقى وراءه، فالطالب الذى يريد أن يأخذ اختبارات لأغراض تعليمية أو للتوجيه المهنى قد يكتشف الفاحص المحنك أن وراء هذه الرغبة الظاهرة مشكلات نفسية واجتماعية هى التى حركت هذا الطالب إلى طلب أخذ الاختبار وغالبا ما يكون هذا الإكتشاف بداية مرحلة من العلاج أو الإرشاد النفسى قد يطلب بعده معلومات من اختبارات، وأنه عليه أن يحسن استخدام هذه المعلومات بعد أن تحل المشكلات التى كانت محاولات أبعادها وأخفائها تستنفد جهده وطاقته، والتى كانت تعوق خطاه عن السير فى الاتجاه المناسب الذى يرتضيه.

ومن الممكن أن يغلب على المفحوص مظهر الخضوع الكامل للفاحص، ويبدى من الرغبة فى التعاون أكثر كثيرا مما يقتضيه الموقف عادة، ومثل هذا إذ يسلم نفسه للفاحص، إنما هو فى الواقع يعبر عن ميله إلى التخلى عن مسئوليته عن نفسه بإلقائها كاملة على الفاحص، بل إن هذا الخضوع قد يخفى وراءه عدوانية أو مقاومة لابد أن يعالجها الفاحص إن كان ذلك فى إمكانه أو يحيلها إلى من يستطيع أن يوفق فيها.

أما عن الفاحص نفسه فإنه ، كما قلنا، يحمل بدوره إلى الموقف الاختبارى تصورات عن نفسه وعن المفحوص ورغبات وشكوكا مما أشرنا إليها، وهذه قد تتوافق

مع تلك التى عند المفحوص بحيث تدعم كل منها الأخرى، وقد تتعارضان وتتضادان، فإذا تصادف أن ذلك المفحوص، الذى يلبس لباس الخضوع ويتخلى عن مسئوليته عن نفسه للفاحص، وقع مع فاحص يميل إلى السيطرة على الناس فإن هذا الفاحص، فى أغلب الظن، لن يستطيع أن يكتشف أن هذه السلبية أو الخضوع ليس إلا حائلا يقيمه المفحوص ليحول بين المختبر وبين الوصول إلى أعماق لا يريد له أن يصل إليها، أما إذا تصادف وكان الفاحص من النوع الذى لا يرحب بأن تحال عليه تلك المسئولية ربما يسبب شك فى نفسه فإنه لن يشعر بالارتياح فى التعامل مع المفحوص. وقد يلجأ إلى الإفراط فى استخدام الاختبارات وما تسفر عنه من نتائج لأن كثرة البيانات تعطيه إحساساً أكبر من الأمن مع مثل ذلك المفحوص.

وقد يكون عند الفاحص دوافع كامنة لايتبينها أثناء قيامه بعملية الاختبار نفسها أو وراء اشتغاله بالاختبار بصفة عامة، فقد يكون اختياره لهذا النوع من العمل إشباعا لحاجات معينة كأن يستشعر الكفاءة الاجتماعية من خلال التفاعل الاجتماعى مع الآخرين وراء أستار الاختبارات ورسمياتها لأنه يحس بعدم القدرة الاجتماعية فى مواقفها الطبيعية، أو قد يكون كمن يقاوم شكوكا أو يجيب عن تساؤلات عن نفسه بمقارنة نتائج مفحوصية بنفسه، أو قد يكون فى تعامله مع المفحوصين ساعيا وراء السند العاطفى وذلك بأن يبدى رقة زائدة معهم ويكون عظيم التسامح والتساهل معهم، مما قد يجعله يغفل جوانب ضعف فى نتائج الاختبار قد يكون لها مغزاها بالنسبة له وللمفحوص على السواء (كالنقص اللغوى). على أن هذا ليس معناه أن يتجنب الفاحص أى تعبير انفعالى مع المفحوص، بل على العكس كما سبق أن أكدنا أنه لا يصح أن يصطنع البرود أو الجمود أو الرسمية المفرطة، وإنما المهم هنا هو أن يكون الفاحص واعيا بدوافعه وحاجاته هو حتى يستطيع أن يبعد آثارها كلما استطاع التأثير فى الموقف الاختبارى الذى هو مسئول عنه، على أن يحرص على إظهار التدعيم العاطفى المناسب الذى ييسر إقامة علاقة سليمة بينه وبين المفحوص ويقلل من مقاومته ويساعده على استدعاء أفضل وأصدق ما عنده فى أدائه فى الاختبار،

(٢) دور اللغة فى تكوين العلاقة بين الفاحص والمفحوص: لعل القارئ يلاحظ أن معظم ما تضمنه موضوع إدراك كل من الفاحص والمفحوص لذاته وللآخر يدخل فى باب التواصل غير اللفظى non-verbal communication بل أن بعضه يدخل

فى مجال «ما لايمكن التلفظ به» unverbalized ونشير هنا إلى الدور المباشر للغة فى تكوين العلاقة بين الفاحص والمفحوص.

ولعل من الأمور الواضحة بذاتها أنه إذا لم يفهم كل من الفاحص والمفحوص لغة الآخر فإن الموقف الاختبارى كله يدخل فى باب العبث أو منطق اللامعقول. ولا نبالغ إذا قلنا أنه يكون أدخل إلى فن الملهاة، وهذا كله بالطبع قد يكون مأساة بالنسبة للقرارات التى تعتمد على نتائج مثل هذه الاختبارات.

ويصدق هذا القول على الباحث الذى يحمل أداته ويذهب بها إلى إحدى القبائل البدائية المنعزلة ثم يقدم لهم تعليمات وأسئلة بلغة لايفهمها إلا هو، كما يصدق على الباحث الآخر الذى يختار عيناته من أعمار معينة (أطفال صغار مثلا) أومستويات اقتصادية اجتماعية معينة ثم يطبق عليهم اختبارات لايكاد يفهم تعليماتها أو أسئلتها سواه، بل قد لايفهمها هو لأنها مترجمة ترجمة شبه حرفية عن لغة أخرى، نقول هذا القول لأنه يوجد فى المكتبة العربية فى الوقت الحاضر طوفان من الاختبارات لم يبذل فيها أصحابها جهدا إلا محض الترجمة ـ التى قد تكون ركيكة ـ إلى اللغة العربية، ونكاد نقول أن بعض هذه الاختبارات لايكاد يفهمها لغويا أصحابها أنفسهم.

وبالطبع توجد لغة الإيماءات والإشارات التى نجح علماء النفس فى استخدامها مع العاجزين ومع الذين لايتحدثون اللغة القومية ومع الأطفال الذين لم يكتمل نموهم اللغوى بعد، إلا أن لهذه اللغة أصولها التى يجب إتقانها والتمكن منها، وإلا أدخلتنا هى الأخرى فى فن العبث أو منطق اللامعقول.

وقد أثبتت البحوث التى أجريت فى الولايات المتحدة (٢٣ : ٦٦ – ٦٧) ـ وهى مجتمع متعدد اللغات ـ أن الأطفال الذين يتقنون لغة معينة ـ كاللغة الإسبانية ـ يجب اختبارهم بهذه اللغة دون اللغة الإنجليزية، إلا إذا كان الهدف من الاختبار قياس مدى اتقانهم اللغة الإنجليزية. وبالطبع فإننا فى اختبار الأطفال يجب أن تستخدم لغة منتقاة من قاموسهم اللغوى، ومن الأفضل بالطبع ـ فى هذه الحالة وأمثالها كاختبار الأميين ـ استخدام اللغة العامية سواء فى إعطاء التعليمات أو فى صياغة أسئلة الاختبار، وفى هذا الصدد لابد للمتخصصين فى مجال القياس النفسى أن يكون على صلة وثيقة بالباحثين فى مجال علم اللغة والمتخصصين فى اللهجات المحلية، وأن يوفر لنفسه

معاجم هذه اللهجات - إن وجدت - ومن المستحسن فى هذه الحالات أن يكون الفاحص الذى يطبق الاختبار على هذه العينات من المتحدثين الأصليين بهذه اللهجة، حتى لاينشأ عن سوء الفهم أو سوء الإستخدام عائقا فى فهم المفحوصين للاختبار.

إلا أن اللغة لا تعمل فى جو اجتماعى فارغ أو أجوف، فلابد فى جميع الأحوال من الدفء العاطفى وجو الصداقة الذى يوفره الفاحص للمفحوص فى موقف الاختبار، وقد أثبتت البحوث التى أجريت على تحليل محتوى المقابلات الشخصية أنه حين تجرى مقابلات رسمية مع أطفال من سن ٨ سنوات يعيشون فى حى الزنوج بمدينة نيويورك (حى هارلم) لم ينتج هؤلاء الأطفال - إلا فيما ندر - استجابات تزيد عن كلمة واحدة، سواء أجريت هذه المقابلات داخل المدرسة أو داخل المنزل. وكان حين يستخدم الأسلوب غير الرسمى فى المقابلة، كانت تتم فى حضور بعض أصدقاء الطفل والجلوس معهم حيث يجلسون - على الأرض مثلا، ومشاركتهم فى شرب علب العصائر التى يشربونها ثم إجراء المقابلة فى صورة حديث عادى، فإن النتائج تثبت فى هذه الحالة أن الطفل الذى يحكم على لغته من خلال المواقف الرسمية للمقابلة بأنها غير ناضجة يندفع منطلقا فى حديث طلق صريح وغير متحفظ.

ولعل القول الأخير يصدق أكثر ما يصدق على الاختبارات التى تتناول الجوانب الوجدانية والاجتماعية لشخصية المفحوص (الميول والاتجاهات مثلا). ولعل أشهر الأمثلة على ذلك تطبيق مقاييس الاتجاهات التى تتناول المسائل الاجتماعية الخلافية الحساسة - كالاتجاه نحو تحديد النسل فى مصر - ان إجابات المفحوصين على هذه المقاييس تأتى مباشرة وقصيرة وتتفق مع قواعد «المرغوبية الاجتماعية» إذا لجأ الفاحص إلى الأسلوب الرسمى، ولكن إذا تبسط مع المفحوصين، كما يجلس مع الفلاحين على «المصطبة»، ويشرب معهم قهوتهم أو شايهم، فإن هذا اللون من الأحاديث العادية» يطلق الألسنة من عقالها، ويجعل الموقف الاختبارى أكثر طبيعية، وحينئذ تؤدى اللغة دورها الطبيعى فى تكوين العلاقة الطيبة بين الفاحص والمفحوص.

كيف يشوه الفاحصون نتائج الاختبارات؟

يسجل تاريخ العلم أمثلة على خداع الذات الذى قد يقع فيه باحثون وملاحظون مدربون، وكثير من هذه الأخطاء يحقق لهؤلاء الفاحصين نوعا من الخدمة الذاتية، قد

لايكون من بينها المكافآت المادية المباشرة لهم، وإنما قد يكون الأهم من ذلك إشباع حاجات الفاحص إلى تدعيم آرائه ومعتقداته.

وبالطبع فإن تشويه نتائج الاختبارات أثناء عملية التقويم النفسى خطأ بالغ، إن لم يكن خطراً عظيمًا، لنفرض أن أحد الباحثين التربويين صمم طريقة جديدة فى تدريس الفهم اللغوى ويرغب - ولو لاشعوريا - فى أن يثبت كفاءتها وفعاليتها، إنه قد يطبق الصورة (أ) من اختبار الفهم اللغوى قبليا، والصورة (ب) بعديا، وبالطبع فإنه فى هذه الحالة قد يضمن الحصول على نتائج إيجابية لصالح طريقة التدريس المقترحة إذا ذكر للمفحوصين أثناء القياس القبلى أن درجات الاختبار ليست لها أهمية ثم قرأ عليهم تعليمات الاختبار بسرعة، وربما يسىء قراءة زمن ساعته فيختصر زمن العمل فى الاختبار، أما إذا استخدم الفاحص عكس ذلك فإنه قد يرفع من درجات مفحوصى المجموعة الضابطة أثناء القياس القبلى، أودرجات مفحوصى المجموعة التجريبية أثناء القياس البعدى، وفى الحالتين يحصل على نتائج مختلفة قد يساء فهمها وتأويلها.

وبعض أخطاء الفاحصين قد تكون مقصودة - وهذه يمكن أن تكون من قبيل الجرائم العلمية التى يذكر تاريخ العلم قليلا منها وصل بعضها بالباحثين إلى حد تزييف الحقائق وفبركة النتائج (١٠)، أما «الأخطاء التى تحركها الدوافع» motivated error فقد تكون غير مقصودة، فأى إجراءا مقنن لتطبيق الاختبار، مهما بلغت دقته لايمنع حدوث بعض الآثار غير المقصودة فى درجات الاختبار حين يتوقع المعلم مثلا من تلميذه «النجيب» أن يؤدى أداء جيدا فيه، وحين يضغط تحيز الفاحص عليه ضغطا لا شعوريا فيظهر الرغبة فى أن يحرز المفحوص من جنسه أو دينه أو بيئته أو طبقته الاجتماعية تفوقا فى الاختبار، ويعود هذا مرة أخرى إلى حاجة المفحوص إلى معرفة «ذاته» وفهم «شخصيته» من حيث ميوله واتجاهاته وقيمه وتفضيلاته وتحيزاته وآرائه وأفكاره، وكلما ازداد استبصار الفاحص بذاته كان أكثر «موضوعية فى مواجهة مواقف الإختبار.

إلا أن هذا الإجراء يحتاج إلى جهد منظم وتدريب شاق، ولعلنا نعيد تنظيم طرق تدريس مقررات الاختبارات النفسية والقياس النفسى والتقويم التربوى والنفسى فى الجامعات بحيث تحتل مسألة تدريب الفاحصين بجوانبها المختلفة التى تناولناها فى

هذا الفصل محور برامج هذه المقررات، ويمكننا أن نستفيد أيضًا من بعض ما يستخدمه الباحثون فى مجالات أخرى - كالطب - حيث يستخدم أسلوب «الفاحص الأعمى المزدوج» double-blind لتجنب الوقوع فى خداع الذات، ومن ذلك أن يعطى للمفحوصين عقار محايد لايختلف على أى نحو ظاهرى عن الدواء الفعلى بحيث لايستطيع الطبيب أو المريض أو مساعد المعمل التمييز بين المجموعتين، ولا يتم هذا التمييز إلا أثناء تحليل النتائج تبعا لنظام ترميزهما، وبالطبع فإن هذا الإجراء لايزال صعب التنفيذ فى مجال البحوث النفسية والتربوية، والبديل هو ما يقترحه بعض الثقاة فى الميدان مثل كرونباك (٢٤ : ٦٨) - من ضرورة أن تطبق الاختبارات فى حالة استخدام المجموعات التجريبية والضابطة بواسطة فاحصين من الخارج لايعرفون مقدما أنواع المجموعات، ثم قبل تصحيح الاختبارات وتفسير نتائجها أن تخلط جميع الأوراق - الاختبارات القبلية والبعدية للمجموعات التجريبية والضابطة - خلطا عشوائيا، مع الاحتفاظ بمفتاح للتعرف على هذه الأوراق، فيما بعد، لايكون متاحا للملاحظين أو المصححين.

وتوجد عوامل مباشرة تكمن وراء تشويه نتائج الاختبارات، سواء كان مقصودا أو غير مقصود، ومن ذلك ما يترتب على درجات الاختبارات من نفع أو خسارة يتعرض لهما الفاحص، ومن ذلك المعلم الذى تعتمد ترقيته أو مكافأته على نتائج فصله فقد يؤثر فى أداء تلاميذه فى الاختبارات بالغش المباشر أو التلويح به، ويذكر هورن Horne وجارتى Garty فى بحثهما عام ١٩٨١ أمثلة كثيرة لما يحدث فى هذا الصدد فى المدارس الأمريكية (٢٤ : ٦٨).

الفصل الرابع

ثبات الاختبارات

معنى ثبات الاختبار

يقصد بمصطلح الثبات reliability فى علم القياس النفسى دقة الاختبار فى القياس أو الملاحظة وعدم تناقضه مع نفسه، واتساقه واطراده فيما يزودنا به من معلومات عن سلوك المفحوص، ولابد للباحث أن يحدد درجة الموثوقية فى بيانات الاختبار كما تتمثل فى اتساق القياسات أو الملاحظات التى يحصل عليها . ويمكن التحقق من هذه الموثوقية بصفة عامة بإحدى وسيلتين: إما بتكرار تطبيق الاختبار نفسه على نفس المفحوصين تحت نفس الشروط والظروف، أو باستخدام اختبارات مماثلة.

وبالطبع فإن من حقائق القياس النفسى أن درجات الاختبارات تختلف من موقف لآخر، وقد يرجع ذلك إلى اختلاف درجة الانتباه عند المفحوص ومقدار جهده المبذول من وقت لآخر، وذلك بالنسبة للفترات الزمنية القصيرة، أما إذا طالت الفترة فإن التغيرات فى درجات الاختبارات قد ترجع إلى التدريب أو التعلم أوالتغيرات فى الجوانب الصحية أو فى شخصية المفحوص، فإذا استخدمت اختبارات مختلفة ـ ولكن متماثلة ـ فإن الاختلاف فى الدرجات قد يرجع إلى مدى ألفة المفحوص بالأسئلة فى كل من الحالتين.

إلا أن أخطر العوامل التى قد تؤدى إلى اختلاف الدرجات فى المواقف الاختبارية المختلفة بالنسبة للاختبار الواحد ولنفس المفحوصين، هو عامل الصدفة chance وهو عامل يرجع إلى الظروف العشوائية التى لايتحكم فيها الفاحص مثل تخمين المفحوص للإجابة أو التذبذبات والتقلبات الوقتية فى انتباهه، ويطلق على هذا العامل خطأ القياس error of measuremcnt.

وفى عملية التقويم النفسى لايمكن أن يعتمد قرار الفاحص على أخطاء المصادفة هذه، وخاصة إذا كان القرار مصيريا مثل انتقاء الشخص لمهنة أو دراسة معينة، أو إعطائه شهادة أو إجازة بالكفاءة أوالخبرة، أو إنهاء العلاج بالنسبة لمريض، أوإلحاق طفل متخلف عقليا بإحدى المؤسسات.

ويستخدم مفهوم الثبات بمعناه الواسع ليدل على مدى اعتماد الفروق الفردية فى درجات الاختبار على أخطاء الصدفة المتضمنة فى القياس (والتى تدل على عدم الثبات) أو على الاختلافات الحقيقية الأصيلة فى الصفة أو الخاصية السيكولوجية موضع الاهتمام (والتى تدل على الصدق كما سنبين فى الفصل التالى)، ويمكن صياغة المشكلة صياغة فنية على النحو الآتى:

إذا افترضنا أن الدرجات التى يحصل عليها المفحوصين فى اختبار معين تمثل التباين الكلى total variance يصبح السؤال ما هى النسبة فى هذا التباين الكلى التى يمكن اعتبارها من نوع تباين الخطأ error variance وتلك التى ترجع إلى التباين الحقيقى؟ true variance إن الإجابة على هذا السؤال ليست سهلة كما يبدو لأن ما نعتبره أحيانا من نوع تباين الخطأ فى موقف معين يمكن اعتباره تباينا حقيقيا فى موقف آخر، فمثلا إذا كان اهتمامنا متجها نحو دراسة التقلب الوجدانى فإن التغيرات اليومية فى درجات اختبار يقيس «الانشراح والانقباض» تصبح مرتبطة بغرض الاختبار نفسه وتكون بالتالى جزءا من التباين الحقيقى فى الدرجات، أما إذا كان الاختبار يهدف إلى قياس الخصائص الأكثر استقرارا فى الشخصية كالذكاء أوالميل العلمى تصبح هذه التغيرات اليومية ذاتها من نوع تباين الخطأ.

ويمكن التعبير عن أنواع التباين المشار إليها بالمعادلة الآتية:

التباين الكلى = التباين الحقيقى + تباين الخطأ.

ومن هذه المعادلة نستنتج أنه إذا زاد التباين الخطأ يقل التباين الحقيقى ويؤدى ذلك إلى تذبذب الاختبار فيما يقيس وعدم اتساقه أو بعبارة أخرى عدم ثباته، والعكس صحيح أى إذا زاد التباين الحقيقى يقل تباين الخطأ ويؤدى ذلك إلى مزيد من استقرار الاختبار وثباته.

وعلى ذلك يمكننا أن نعرف التباين الخطأ بأنه يتعلق بالشروط والظروف التى لا تتصل من قريب أو بعيد بغرض الاختبار، ولذلك فحينما يحاول المدرس أو المرشد النفسى أن يقنن الاختبار بأن يحافظ على توحيد الظروف الاختبارية وذلك بضبط مكان الاختباروتعليماته وزمنه وعلاقته بالمختبرين وغيرها من العوامل التى تحدثنا عنها فى الفصل السابق فإنه بذلك يقلل من فرص حدوث تباين الخطأ ويجعل درجة

الاختبار أكثر ثباتا واستقرارا، ومع ذلك فلا يمكن اعتبار الاختبار أداة ثابتة ثابتا كاملا يصل إلى درجة ثبات مقاييس الظواهر الطبيعية (كالمتر للأطوال والكيلو جرام للأوزان) . ولذلك لابد أن يحدد صانعوا الاختبارات مقدار ثباتها فى كراسة التعليمات بحيث يكون دليلا للأخصائيين النفسيين حينما تطبق هذه الاختبارات فى ظروف معيارية مقننة .

والمشكلة الجوهرية فى تحديد ثبات الإجراءات الاختبارية بالنسبة لمجموعة معينة من الأفراد هى فى التمييز الحقيقى والتباين الذى يرجع إلى الخطأ، وعلاقة ذلك بنوع الاستنتاج الذى يرغب الباحث النفسى فى التوصل إليه من درجات الاختبار، وبعدما يتم هذا التمييز بين نوعى التباين يبدأ الباحث فى تصميم سلسلة من العمليات الأمبريقية (التجريبية) والإحصائية تفيد فى الوصول إلى أفضل تقدير لنسبة كل من نوعى التباين.

مصادر التباين :

١ – **الخصائص المستمرة والعامة فى الفرد:** فالتباين فى الدرجات يرجع إلى ثورنديك أول من أعد تصنيفا شاملا لهذه المصادر عام ١٩٥١ فى فصل عظيم كتبه عن «الثبات» فى كتاب لنديكويست الشهير (٢٩)، وقد عدل ستانلى J. C. Stanley هذا التصنيف فى فصل جديد بنفس العنوان فى الطبعة الثانية من هذا الكتاب والتى حررها روبرت ثورنديك (٣٤)، وتشمل هذه القائمة المعدلة ما يلى (٣٤ : ٣٦٤):

١ – **الخصائص المستمرة والعامة فى الفرد:** فالتباين فى درجات الأفراد ينشأ فى العادة نتيجة لما يظهره الأفراد المختلفون من مقادير مختلفة فى سمات أو خصائص معينة ذات طبيعة مستمرة وعامة، ومن ذلك الاختلاف فى مستوى قدرة معينة لها دورها فى الأداء فى عدد كبير من الاختبارات ويؤدى هذا الاختلاف إلى فروق فردية ظاهرة ومن أمثلة ذلك القدرة على الاستدلال أو السهولة أو الفهم اللغوى. وعلى هذا توجد فى الأفراد. اختلافات بقدر ما وجد من سمات أو قدرات، وهى كثيرة العدد بالغة التنوع لايتسع المقام لحصرها، ويمكن الرجوع إليها فى المراجع المتخصصة (١) . ويمكن القول أن أى أداء اختبارى يعتمد جزئيا على القدرات

والسمات التى تتداخل بالفعل فى عدد من الاختبارات النفسية.

ولعل القارئ يلاحظ أن هذا المصدر من مصادرالتباين يمثل خصائص مستمرة ومستقرة لدى الفرد، ويؤدى بالضرورة إلى فروق فردية مستقرة وحقيقية فى الأداء الاختبارى، وهو نوع من التباين المنتظم على نحو يعين الباحث على تناوله فى أى إجراء يتخذه عند تقدير الثبات.

ويضيف ستانلى إلى فئة الخصائص المستمرة والعامة مصدرين آخرين للتباين هما: القدرة العامة على فهم التعليمات والخبرة بأخذ الاختبارات test-wiseness وهما عاملان يكادان يؤثران فى كل درجة اختبارية سواء أراد الفاحص أو لم يرد، ويلعب عامل فهم التعليمات دوراً مهماً على وجه الخصوص حين يكون الاختبار جديدا، وكانت تعليماته معقدة، أما العامل الثانى فيتمثل فيما يظهره المفحوص، «الخبير بالاختبار» من خصائص تختلف عن المفحوص الأقل خبرة، فالأول يكون عادة أقل انفعالية، وأمهر فى التخمين، وأكثر وعيا بالمنبهات الثانوية والخارجية التى تفيد فى الإجابة على أسئلة الاختبارات، ولعلك تلاحظ أن وجود تباين ناتج عن هذين المصدرين يعد من الأمور غير المرغوب فيها فى ضوء الغرض من الاختبار إلا إذا ارتبط بقدرة عقلية أساسية، ويمكن التقليل من أثرهما بصفة عامة عن طريق تحسين تعليمات الاختبار من ناحية وتقديم قدر من التدريب على أسئلة الاختبار من ناحية أخرى.

وتمثل هذه الفئة من مصادر التباين تحديا حقيقيا لمؤلفى الاختبارات لأنها تتطلب التحكم، إلا إذا كان وجودها مطلوبا ومرغوبا فيه وفى هذه الحالة تصبح من نوع التباين الحقيقى، ويتطلب هذا من الباحث النفسى أن يقدم لمفحوصيه تعليمات على أعلى درجة من الوضوح وبأقل قدر من المنبهات الثانوية.

٢ - **الخصائص المستمرة والخاصة فى الفرد:** قد ينشأ التباين فى الدرجات الاختبارية نتيجة لخصائص معينة فى الأفراد تتسم بالاستمرار من ناحية إلا أنها تكون على درجة من الخصوصية للمجال النوعى الذى نختبره، ومن أمثلة ذلك ما نجده فى اختبارات التهجى أو القسمة المطولة دون سواها من الاختبارات، فبالإضافة إلى التباين الذى يميز الأداء العام فى التهجى أو الحساب يوجد نوع آخر من

التباين يرتبط بصورة معينة من صور الاختبار أو طريقة معينة فى هذا الاختبار كما هو الحال مثلا فى سلوك التهجى حين يكون على المفحوص أن يتهجى كلمات معينة شفويا أو يكتبها صحيحة فى الإملاء أو يتعرف على الأخطاء فى بعض الكلمات المطبوعة. واختبار القسمة المطولة التى يتطلب إنتاج الاستجابة أو اختيارالإجابة من عدة بدائل (الاختيار من متعدد)، وقد ينشأ عن ذلك نوع من التباين بالطبع، ويوجد مصدر آخر للتباين ينتمى إلى هذه الفئة هو طبيعة عينة الأسئلة أوالمفردات التى يتألف منها الاختبار وينشأ عن الفروق الفردية فى تفاصيل المعارف والمهارات التى لدى الأفراد المختلفين، وعلى هذا لو أعد الباحث اختبارين من عينتين من الأسئلة اختيرتا مستقلتين من نفس الأصل الإحصائى للأسئلة أوالمفردات فقد لا يحصل الفرد الواحد على الدرجة نفسها فيهما بسبب التنوع بين الأفراد فى المهارات والمعلومات اللازمة للإجابة على أسئلة كل من الاختبارين. وفى هذه الحالة تنشأ الحاجة إلى المقارنة بين الأفراد فى اختبارات متماثلة فيما تقيس وتختلف فى أنواع المفردات أو طريقة العرض للتحكم فى هذا المصدر باعتباره ينتمى إلى تباين الخطأ.

٣ - الخصائص العامة المؤقتة فى الفرد أوالموقف الاختبارى: من أمثلة هذه الخصائص الصحة الجسمية ومقدار النوم الذى حصل عليه الفرد فى الليلة السابقة على الاختبار، والتعب، وحالة الدافعية، والتوتر الانفعالى وغيرها من العوامل الداخلية فى الفرد والتى تؤثر على أدائه فى الاختبار، بالإضافة إلى العوامل الخارجية فى الموقف الاختبارى كدرجة حرارة الجو، والإضاءة والتهوية وغيرها، ومن المعلوم أن أداء الفرد فى الاختبارات المختلفة يتأثر بهذه العوامل بدرجات مختلفة ، إلا أن جميع الاختبارات تتأثر بها على نحو أو آخر، كما تختلف هذه العوامل فى درجة استمرارها وعموميتها، فبعضها قد يتغير من يوم لآخر، وبعضها الآخر من ساعة لأخرى، وبعضها الثالث من لحظة لأخرى، ويمكن لنا أن نتصور هذه العوامل على أنها تميز الفرد فى جلسة اختبار معينة دون سواها، ويحتاج الباحث النفسى فى مثل هذه الأحوال إلى التحقق من اتساق الأداء فى الاختبار سواء فى لحظة معينة هى لحظة تطبيق الاختبار، أو من فترة زمنية إلى أخرى (بعد عدة أيام أو عدة أسابيع) كما أن على الباحث أن يحدد ما إذا كانت هذه الخصائص المؤقتة من المصادر المنتظمة الحقيقية للتباين أم هى من مصادر تباين الخطأ.

٤ - **الخصائص الخاصة المؤقتة فى الفرد:** ويشمل هذه الفئة العوامل ذات النطاق الضيق سواء من حيث الزمن أو المدى إذا قورنت بعوامل الفئة السابقة، وبعضها يرتبط بالاختبار ككل، وبعضها يرتبط بأسئلته، ومن أمثلة النوع الأول ما يتصل بصعوبة التعليمات فى حالة الاختبارات الجديدة التى قد تنشأ عنها فروق فردية فى فهم طبيعة المهام التى يتألف منها الاختبار. وقد يؤدى ذلك إلى تباين عشوائى ينتمى إلى تباين الخطأ. وكذلك قد تتطلب الإجابة على الاختبار استراتيجيات مختلفة يستخدمها الأفراد المختلفون، بعضها قد يكون أكفأ من البعض الآخر، وحينئذ ينشأ مصدر جديد لتباين الخطأ، وبالإضافة إلى ذلك فإن الأداء فى بعض الاختبارات - وخاصة اختبارات المهارات الحركية - قد يكون عرضة لتحسن نتيجة للممارسة، ويؤثر فى ذلك مستوى الممارسة لدى الفرد عند بدء الاختبار. كما قد يوجد مجموعة من العوامل المتنوعة التى يشملها مصطلح «التأهب لأداء الاختبار»، مثل التركيز على السرعة أو الدقة فى اختبارات القدرات الموقوتة، أو الوعى ببعض المنبهات فى حالة الاختبارات الإدراكية، أو الحالة المزاجية لدى الفرد فى حالة مقاييس الاتجاهات أو قوائم الميول، وهذه العوامل - على الرغم من أهميتها - يمكن تجاهلها فى الاختبارات البسيطة والمقننة، إلا أنها تلعب دوراً أكبر فى الاختبارات الموقوتة، وفى الاختبارات التى تتكون من أنواع جديدة من الأسئلة أوالمفردات، وفى المقاييس التى تتطلب مقدارا كبيرا من التفسير الاستبطانى، وفى الأداء الاختبارى الذى لم يتجاوز المرحلة التكوينية (أى لم يصل فيه المفحوص إلى إجابة نهايته، وإنما تكون أجابته على هيئة مسودة مبدئية).

وتوجد مجموعة أخرى من العوامل تنتمى إلى هذه الفئة وتخص المفردات أو الأسئلة ذاتها التى يتكون منها الاختبار، أو بضع لحظات من الزمن الكلى للاختبار وتتمثل فى التذبذبات الوقتية فى الذاكرة أو الانتباه، أو التعويق المؤقت للأفراد، أو التنوع الدورى فى مقدار الجهد المبذول، وهى جميعا تنتمى إلى تباين الخطأ.

٥ - **العوامل المنتظمة أوالعشوائية التى تؤثر فى تطبيق الاختبار أو تصحيحه:** قد ينشأ التباين من مصادر أخرى غير المفحوص ذاته وتنشأ عن ظروف إعطاء الاختبار وتصحيحه، ومن ذلك زمن الاختبار وتعليماته ومقدار الضوضاء وتشتت الانتباه وفشل الإضاءة ونقص مواد الاختبار (كالأقلام) وغيرها، وهى جميعا

تختلف من شخص لآخر أو من مجموعة لأخرى وهذا النوع من الاختلاف ينشأ على وجه الخصوص حين يطبق الاختبار الواحد فاحصون مختلفون، أو فى الاختبارات التى تحتاج إلى توقيت دقيق أو تلك التى تتطلب تعليمات معقدة للغاية، كما قد ينشأ هذا المصدر للتباين عند تصحيح أوتقدير الدرجة فى الاختبار وخاصة حين يعتمد ذلك على حكم الفاحص أو الملاحظ أوالمصحح، ومن ذلك إعطاء الدرجة أو الحكم فى اختبارات المقال، ومقاييس التقدير أوالأساليب الاسقاطية أو فى أى اختبار يتطلب قدرا من التفسير أو الحكم من جانب الفاحص - إننا فى هذه الحالة قد تجدنا إزاء نوع من التباين يرجع فى أصله إلى اختلاف الفاحصين أو الملاحظين أو المصححين، أو إلى التغيرات التى تطرأ على الفاحص الواحد من وقت لآخر، أو إلى تناقض لايخضع للتحكم أو التنبؤ داخل الفاحص أو المصحح نفسه، وهذه العوامل جميعا تنتمى إلى تباين الخطأ.

٦ - **العوامل التى ترجع إلى الصدفة البحتة:** يوجد نوع آخر من التباين لا تستطيع أن تعود به إلى عامل بذاته (أو عوامل بذاتها) ولذلك يمكن أن نعتبر عوامله من النوع العشوائى أوالذى يرجع إلى الصدفة، ومن ذلك تخمين المفحوص للإجابة فى الاختبار حيث يلعب «الحظ» دوراً واضحاً فى انتقاء الإجابة (والتى قد تكون صحيحة)، كما قد يرجع إلى عوامل أخرى كثيرة غامضة تلعب دورها فيتشتت انتباه المفحوص فى لحظة معينة وهذه العوامل تنتمى إلى تباين الخطأ بمعناه المباشر.

طرق حساب الثبات:

توجد بالطبع أنواع من ثبات الاختبار بعدد مصادر التباين التى أشرنا إليها، أى بعدد الظروف التى يمكن أن تؤثر فى درجات هذه الاختبارات ولا تكون كلها أو بعضها متعلقة بغرض الاختبار، وتوجد طرق تستخدم فى حساب الثبات يشير كل منها إلى المصادر التى يمكن أن يرجع إليها تباين الخطأ، وهذه الطرق ليست بدائل يحل بعضها محل بعض، وإنما لكل منها استخدامه الخاص حسب مصدر تباين الخطأ الذى نهتم به.

١ - إعادة الاختبار:

أشرنا إلى أن من المصادر الأساسية لتباين الخطأ ما يتمثل فى التغيرات العشوائية التى تنتاب الأفراد من جلسة اختبار إلى أخرى، وقد تنتج هذه التقلبات عن ظروف تطرأ على الموقف الاختبارى ولايمكن التحكم فيها مثل التغيرات الفجائية فى الجو أو

حدوث ضوضاء مفاجئة (وغيرها من مشتتات الانتباه). وقد تنتج عن تغيرات فى ظروف المفحوص نفسه كما تتمثل فى حالات المرض والتعب والاضطراب الانفعالى والقلق وغيرها، وهذه العوامل تدل على المدى الذى تتأثر به درجات الاختبار بالتقلبات اليومية العشوائية فى المفحوص أو فى ظروف الموقف الاختبارى.

ومن الواضح أن درجة الاستقرار الوقتى Temporal Stability تعتمد إلى حد ما على الفترة الزمنية التى تنقضى على تطبيق الاختبار والتى نقدر بعدها ثبات الاختبار. وتوجد أمثلة عديدة لاختبارات لها ثبات مرتفع إذا قدر بعدها ثبات الاختبار، وتوجد أمثلة عديدة لاختبارات لها ثبات مرتفع إذا قدر هذا الثبات بعد فترات زمنية قصيرة (أيام قلائل أو أسابيع قلائل) إلا أن ثباتها ينخفض انخفاضا شديدا عندما تمتد هذه الفترات إلى عدة سنوات، وكثير من اختبارات ذكاء أطفال ما قبل المدرسة لها ثبات مرتفع خلال فترة ما قبل المدرسة ولكنها تفشل فى التنبؤ بذكاء الأطفال فى مرحلة الطفولة المتأخرة أوالمراهقة أوالنضج بسبب ما يطرأ على الطفل من تغيرات ارتقائية فى النواحى العقلية.

ولاشك فى أن التذبذبات ذات المدى الزمنى القصير يمكن اعتبارها من نوع تباين الخطأ فى درجة الاختبار، ولذلك ففى الممارسة العملية عندما تقوم بتقدير مثل هذا النوع من الثبات بإعادة الاختبار لابد من أن تكون الفترة الزمنية بين مرتى التطبيق قصيرة، وتكون أقصر عند اختبارنا للأطفال الصغار منها عند اختبارنا للكبار بسبب سرعة معدل النمو فى الطفولة عنه فى مراحل النمو الأخرى، وبالنسبة لأى نوع من المفحوصين فإن قواعد الخبرة تؤكد أنه يجب ألا تزيد فترة ما بين الاختبارين عن ستة أشهر.

هذا التحفظ مهم للغاية، لأننا نعلم جميعا أن السلوك الإنسانى فى حالة تغير مستمر، فإذا كانت هذه التغيرات تحدث فى الأداء الاختبارى للمفحوصين بعد انقضاء فترات زمنية طويلة فإنها يمكن أن تعد ضربا من النمو أو التدهور فى السلوك لامجرد تغيرات عشوائية، بل قد تنتمى هذه التغيرات إلى ميدان من السلوك أوسع نطاقا مما يتناوله الأداء الاختبارى ذاته، فالمفحوص قد يتغير أداؤه فى اختبارات الاستعداد المدرسى أو الفهم الميكانيكى أوالحكم الفنى بعد عشر سنوات بسبب تعرضه لخبرات

غير عادية خلال هذه الفترة، بعبارة أخرى يمكن للفرد أن يتغير مركزه النسبى فى الاختبار بعد هذه الفترة صعوداً أوهبوطاً بالنسبة لأقرانه من نفس العمر بسبب ظروف خاصة تعود إلى المنزل أو المدرسة أو البيئة الاجتماعية أو بسبب المرض المزمن أو الاضطراب الانفعالى المستمر.

ويجب أن نشير إلى أن العمليات النفسية ذاتها قد تختلف فيما بينها فى مدى ما تظهره من تذبذبات وتقلبات وقتية، فمثلا قد يتأثر اتزان الحركات الدقيقة للأصابع بأى تغير ولو كان طفيفا فى ظروف المفحوص أو فى ظروف الموقف الاختبارى بينما نجد الاستدلال لا يسهل التأثير فيه بالمؤثرات العارضة، ولذلك إذا أردنا أن نحصل على تقدير دقيق لثبات اختبار اتزان الأصابع يجب إعادة تطبيقه مرات عدة فى أيام متوالية، بينما قد تكفى جلسة اختبار واحدة لقياس الاستدلال دون حاجة إلى إعادة التطبيق ومعنى ذلك أنه من الضرورى فهم طبيعة السلوك الذى يقيسه الاختبار قبل اختيار أنسب الطرق لحساب ثباته.

ويقدر معامل الثبات الذى يتأثر بتباين الخطأ الناتج عن التقلبات الوقتية بطريقة إعادة الاختبار test-retest . ويطبق الاختبار فى المرة الثانية، بصورته ذاتها على نفس الأفراد، وقد يتطلب ذلك إعادة التطبيق عدة مرات فى مناسبات مختلفة، ويحسب معامل الارتباط بين درجات المفحوصين فى مرتى التطبيق أو متوسط معاملات الارتباط إذا أعيد تطبيق الاختبار عدة مرات، ولعلك تلاحظ أن درجات الأفراد قد تختلف فى المرة (أو المرات) المختلفة، وهذا لايهم كثيرا إنما المهم أن أوضاعهم النسبية لا تتغير كثيرا، ومعامل الارتباط فى جوهره هو طريقة تحديد الوضع النسبى ويسمى المعامل الناتج معامل الاستقرار coefficient of stability.

ورغم بساطة هذه الطريقة فإنها تتضمن بعض المشكلات، منها مثلا أن الممارسة قد تنتج قدرا من التحسن فى درجات إعادة الاختبار، ومنها أيضاً أنه لو كانت الفترة بين موقف الاختبار وإعادته قصيرة فإن المفحوص قد يستدعى بعض استجاباته السابقة، أو بعبارة أخرى قد نحصل على نفس النمط من الإجابة من حيث الوضع النسبى للصواب والخطأ ولكن ليس بسبب العملية النفسية موضع القياس، وبالتالى لاتكون الدرجات التى نحصل عليها من موقفى الاختبار مستقلة بعضها عن

بعض، كما أن معامل الارتباط بين مجموعتى النتائج قد يكون مرتفعا ارتفاعا زائفا، وكذلك من مشكلات طريقة إعادة الاختبار أن طبيعة الاختبار ذاتها تتغير بسبب الإعادة وخاصة إذا كانت أسئلة الاختبار تتطلب استخدام بعض العمليات العليا كالاستدلال أو الابتكار. ففى مثل هذه الأسئلة نجد أن المفحوص إذا استطاع اكتشاف العلاقة أو أعطى استجابة جديدة أو توصل إلى حل المشكلة، فإنه يمكن أن يستعيد أجابته ذاتها فى المستقبل اعتمادا على الذاكرة دون التفكير ، ومعنى ذلك أن طبيعة الأداء فى الموقفين تكون مختلفة، ومع ذلك فإن معامل الارتباط بين مرتى التطبيق قد يكون مرتفعا ارتفاعا زائفا، ومع ذلك لايدل على ثبات حقيقى، وعلى ذلك فإن هذه الطريقة لا تصلح إلا مع الاختبارات التى لا تتأثر تأثيرا كبيرا بالإعادة أو التكرار مثل اختبارات التمييز الحسى السمعى أو البصرى، أما بالنسبة لمعظم الاختبارات النفسية فإن طريقة إعادة الاختبار بالرغم من شيوعها فى الاستخدام قد لاتكون ملائمة لها.

٢ – الصور المتكافئة أو البديلة :

من المألوف لدى كثير من الطلاب أن يدخلوا الامتحان «فيسعدهم الحظ» بعدد كبير من الأسئلة يتناول الموضوعات التى اهتموا بدراستها ومذاكرتها أكثر من غيرها، وقد يحدث العكس فى كثير من الأحيان هذا الموقف يدل على مصدر آخر من مصادر تباين الخطأ فى درجات الإختبار وهو مدى اعتماد درجة المفحوص فى الاختبار على عوامل تتعلق بطريقة اختيار الأسئلة، فإذا افترضنا أن قام باحث آخر بإعداد اختبار تتوافر فيه نفس مواصفات الاختبار الأول مع اختلاف فى المحتوى فإلى أى حد تختلف درجة المفحوص فى الاختبارين؟ الواقع أن الفرق فى الدرجة التى يحصل عليها نفس الأفراد فى هذين الاختبارين هو نوع الثبات التى يتأثر بعينة الأسئلة التى يتكون منها كل منهما. فمثلا إذا افترضنا أن باحثا قام بإعداد اختبار لقياس المعلومات الميكانيكية يتكون من ٥٠ سؤالا، ثم قام باحث آخر بإعداد اختبار آخر مكافئ له (أى يتكون من ٥٠ سؤالا مختلفا تقيس نفس العملية النفسية) ، ولنفرض أن أسئلة كل من الاختبارين قد صيغت بنفس الدرجة من المهارة بحيث يتناول نفس المستوى من الصعوبة، فإنه تبعا لعوامل الصدفة البحتة الناجمة عن الفروق الفردية فى الخبرات السابقة للمفحوصين قد تختلف الصعوبة النسبية لمجموعتى الأسئلة من شخص لآخر،

وعلى ذلك فإن الاختبار الأول قد يحتوى على عدد كبير من الأسئلة لم يألفه المفحوص (أ) بينما تكون ألفته بأسئلة الاختبار الثانى أكبر، وقد يحدث العكس بالنسبة للمفحوص (ب).

فإذا افترضنا أن هذين المفحوصين متساويان تقريبا فى معلماتهم الميكانيكية (أى فى درجاتهم الحقيقية أو التباين الحقيقى) فإن المفحوص (ب) سوف يتفوق على المفحوص (أ) فى الاختبار الأول، بينما سوف يتفوق (أ) على (ب) فى الاختبار الثانى، ومعنى ذلك أن الموقف النسبى للمفحوصين قد يتغير بل قد ينعكس ويمثل هذا نوعا من تباين الخطأ يرجع إلى الصدفة الناتجة من طبيعة محتوى كل من صورتى الاختبار.

ويقدر معامل الثبات فى هذه الحالة بتطبيق صورتى الاختبار فى جلسة واحدة أو فى جلستين مختلفتين على نفس الأفراد، ثم يحسب معامل الارتباط من الدرجات فى الاختبارين ويسمى معامل الثبات فى هذه الحالة بمعامل التكافؤ Coefficient of equivalence، والمهم فى جميع الحالات أن تكون صور الاختبار متكافئة فعلا، أى تقيسان نفس العمليات النفسية بنفس المستوى من الصعوبة أو الألفة، وأن تؤثر هذه العمليات النفسية فى الأداء فى الاختبار بنفس الطريقة، ولكن، كما سبق أو أوضحنا، قد تؤثر الصفات الخاصة تأثيرا مختلفا فى الاختبارين ومن ذلك المعلومات الخاصة التى قد تساعد المفحوص فى الإجابة على سؤال معين من أسئلة أحد الاختبارين ولكنها لا تساعده فى الاختبار الآخر.

وتباين الخطأ المتضمن هنا يدل على التغيرات فى الأداء من مجموعة معينة من الأسئلة إلى مجموعة أخرى، أى الخطأ الذى يرجع إلى عينة الأسئلة ولكنه لايتضمن تباين الخطأ الذى يعود إلى العوامل الوقتية، والذى سبق أن تحدثنا عنه، والذى يتمثل فى التغيرات فى الأداء من موقف اختبارى إلى آخر، ويمكن الجمع بين هذين النوعين من تباين الخطأ إذا قمنا بتطبيق الصورتين المتكافئتين من الاختبار فى مناسبتين مختلفتين (أى لانطبقهما فى الوقت نفسه)، ويمكن بهذه الطريقة تفادى بعض مشكلات طريقة إعادة الاختبار ذاتها، وفى هذه الحالة يدل معامل الثبات على كل من الاستقرار والتكافؤ معا، ويسمى بهذا الاسم.

وعند إعداد الصور المتكافئة equivalent أو البديلة alternate من الاختبارات يجب أن نتأكد كما قلنا من أن هذه الصور فعلا متكافئة، والصور المتكافئة هى فى جوهرها اختبارات تم إعداد كل منها على حدة وبطريقة مستقلة بحيث تتوافر فيها نفس المواصفات، أى أن تحتوى على العدد نفسه من الأسئلة، وأن تكون صياغة هذه الأسئلة متماثلة، وأن تتضمن الأسئلة محتوى متشابها، وأن تتعادل الأسئلة فى مستويات الصعوبة، وأن تتفق الاختبارات فى جميع المظاهر الأخرى مثل التعليمات والزمن والأمثلة التوضيحية والشكل العام. ولايمكن أن تعزى الفروق فى الدرجات من اختبار إلى آخر إلى تباين الخطأ الذى يرجع إلى عينة محتوى الاختبار إلا إذا كانت الصورتان متكافئتين أومتعادلتين حقا.

ومن المؤكد أن حاجتنا إلى الصور المتكافئة من الاختبارات تتعدى مجرد تحديد ثبات الاختبار، فالصور البديلة تفيد كثيرا فى الدراسات التى نجريها حول آثار عامل تجريبى معين فى الأداء الاختبارى وخاصة إذا تطلب الأمر قياسا قبليا وبعديا، كما يفيد استخدام الصور البديلة فى إعطاء وسيلة لخفض احتمال التدريب أوالغش، إلا أن أهم صعوبات هذه الطريقة تتمثل فى عدم إمكان إعداد صور متكافئة لكثير من الاختبارات لأن ذلك يضاعف الجهد والوقت والتكلفة المبذولة فى بناء الاختبار، ومن المعروف أن إعداد اختبار نفسى واحد يتطلب جهدا عظيماً، فما بالك بإعداد أكثر من اختبار لقياس نفس السمة.

٣ - التجزئة النصفية:

يدل تجانس أسئلة الاختبار على مدى اتساق واطراد أداء المفحوص على جميع الأسئلة التى يتكون منها الاختبار، فإذا كان لدينا اختباران تتكون أسئلة الأول منهما من عمليات ضرب فقط بينما تتكون أسئلة الاختبار الثانى من العمليات الحسابية الأساسية الأربعة فإن الأداء فى الاختبار الأول سوف يكون أكثر اتساقا واطرادا بين أسئلته من الاختبار الثانى، لأنه فى الاختبار الثانى (وهو اختبار غير متجانس الأسئلة) قد يحسن بعض المفحوصين الأداء فى عمليات الطرح مثلا بينما يسيئون أداء باقى العمليات، فإذا زاد عدم تجانس الأسئلة كأن يتكون الاختبار مثلا من ٤٠ سؤالا يقيس كل عشرة منها عمليات نفسية مختلفة كالعلاقات المكانية والاستدلال الحسابى

والمفردات والسرعة الإدراكية فإنه قد تكون العلاقة ضئيلة للغاية (إن كانت هناك علاقة) بين أداء المفحوص على هذه الأنماط الأربعة المختلفة من الأسئلة.

من ذلك نتبين أن درجة الشخص فى اختبار ما تكون أكثر وضوحا عندما نحصل عليها من اختبارات متجانسة البنية نسبياً، وتبرز مشكلة وضوح معنى الدرجة الاختبارية إذا عدنا مرة أخرى إلى المثال الأخير، فإذا حصل المفحوصان (أ) و (ب) على درجة مقدارها ٢٠ فى هذا الاختبار فهل يمكننا أن نستنتج من ذلك أن أداءهما متساو تماما فى هذا الاختبار؟ الإجابة بالنفى دون شك، فقد ينجح المفحوص (أ) مثلا فى الإجابة على أسئلة المفردات والسرعة الإدراكية ويفشل تماما فى الإجابة على أسئلة الاستدلال الحسابى والعلاقات المكانية، أما المفحوص (ب) فقد يحصل على درجته من الإجابة على ٥ أسئلة للسرعة الإدراكية و٥ أسئلة أخرى للعلاقات المكانية و١٠ أسئلة للاستدلال الحسابى ويفشل تماما فى الإجابة على أسئلة المفردات، ويوجد عدد كبير من الاحتمالات الأخرى للدرجات التى يحصل عليها المفحوصون من الاختبارات غير المتجانسة فى البنية، ومعنى ذلك أن درجة المفحوص فى مثل هذا الاختبار لها معان مختلفة، بينما يكون معناها أكثر وضوحا فى الاختبارات المتجانسة.

والدرجة فى الاختبار المتجانس الأسئلة (معانى المفردات مثلا) قد لايكون لها أيضًا معنى واحد تماما فى جميع الحالات، وإنما قد تعنى مثلا أن المفحوص أجاب إجابة صحيحة على المفردات العشرين الأولى تقريبا إذا كانت هذه المفردات مرتبة حسب الصعوبة، وربما قد يترك كلمتين أو ثلاث من الكلمات السهلة أو قد يفشل فى الإجابة عليها، وقد ينجح فى الإجابة على كلمة أو كلمتين أصعب وأبعد فى الترتيب ومع ذلك فإن هذه الاختلافات الفردية قليلة وهينة إذا قورنت بتلك التى نلاحظها فى الاختبارات الأقل تجانسا.

وهكذا يعد عدم تجانس أسئلة الاختبار مصدرا لتباين الخطأ، وبالتالى يؤثر فى ثبات الاختبار، ولكن هل يختلف تجانس الأسئلة كمصدر لتباين الخطأ عن المصدر الذى سبق الحديث عنه وهو ملاءمة عينة المحتوى عموما؟ الواقع أن الفرق بين هذين المصدرين يمكن أن يتضح إذا أعطينا مثالا متطرفا، لنفترض أن هناك اختبارا يقيس كل سؤال فيه عملية نفسية مختلفة ولايرتبط أدنى ارتباط بالأسئلة الأخرى التى

يقيس كل منها عمليات أخرى، أننا فى مثل هذه الحالة يمكن أن نعد صورة متكافئة من هذا الاختبار تحتوى على نفس أنواع الأسئلة وتوزيعها على نفس النحو الذى هى عليه فى الصورة الأولى، وعندما نعطى الصورتين المتكافئتين قد تتفق درجاتهما اتفاقا كاملا مما يدل من الوجهة النظرية على معامل ثبات مرتفع لايتأثر تأثرا كبيرا بعينة الأسئلة كمصدر لتباين الخطأ ولكن عندما نهتم بالثبات فى ضوء تجانس الأسئلة نجد أن هذا المعامل يصل إلى الصفر أو يقترب منه، حيث أن إطراد الأداء من سؤال إلى آخر فى إى من صورتى الاختبار قد لا يتجاوز حدود الصدفة.

وللتحقق من أثر عدم تجانس الأسئلة كمصدر لتباين الخطأ فى الاختبار يلجأ الباحث إلى تطبيق صورة واحدة من الاختبار مرة واحدة، ثم يقوم عند تصحيح الإختبار بتجزئته إلى نصفين، وعندئذ يحصل على درجتين لكل مفحوص كل منهما فى كل نصف من نصفى الإختبار يمكنه المقارنة بينهما ويسمي هذه الطريقة التجزئة النصفية split-half، ويحسب الباحث معامل الارتباط بين درجات الاختبار التى حصل عليها المفحوصون فى نصفى الاختبار، ويسمى معامل الثبات الناتج معامل الاتساق الداخلى internal consistency، حيث أن الأمر لايحتاج أكثر من تطبيق صورة واحدة للاختبار مرة واحدة فقط، وعملية التجزئة تتم أثناء تصحيح الاختبار وليس أثناء تطبيقه.

ولعل أخطر المشكلات التى تواجهها طريقة التجزئة النصفية هى كيف نجزيء الاختبار للحصول على أفضل قسمين للمقارنة، أى قسمين متكافئتين، وخاصة إذا علمنا أن معظم الاختبارات لايمكن تجزئتها إلى نصفين صالحين للمقارنة بينهما بمجرد تتبع التسلسل الرقمى للأسئلة، وذلك بسبب الفروق فى طبيعة الأسئلة ومستوى صعوبتها بالإضافة إلى الآثار التراكمية للحماس فى الأداء أو العمل والممارسة والتعب والملل وغيرها من العوامل التى تختلف فى بداية الاختبار عنها فى نهايته، ولذلك فإن إحدى الطرق الدقيقة والموضوعية للحصول على قسمين متناظرين ومتكافئتين فى الاختبار أن يكون الاختبار فى أعداده المبدئى قد تحددت فيه مستويات الصعوبة لكل سؤال وذلك بإيجاد النسب المئوية للأفراد الذين أجابوا عليها إجابة صحيحة، وعندئذ توزع الأسئلة على قسمى الاختبار تبعا لتكافؤ مستوى الصعوبة وتشابه المحتوى، فإن كان ترتيب الأسئلة فى الاختبار متسلسلا تبعا لمستويات السهولة أو الصعوبة فإن

الباحث قد يستخدم طريقة صالحة لمعظم الأغراض وفى الوقت نفسه لا تستغرق وقتا طويلا ولا تتطلب جهداً زائداً وهى تجزئة الاختبار إلى نصفين أحدهما يضمة الأسئلة ذات الأرقام الزوجية والآخر يضم الأسئلة ذات الأرقام الفردية، إلا أننا يجب أن نحذر أنه فى حالات مجموعات الأسئلة التى تتعلق بمشكلة واحدة (مثل الأسئلة التى تتلو فقرة معينة فى اختبار لفهم قراءة) فإن هذه الفقرة بأسئلتها جميعا يجب أن توضع فى أحد قسمى الاختبار من الاختبار، لأننا لو وزعنا مجموعة الأسئلة التى من هذا النوع على نصفى الاختبار فإن التشابه بينهما سوف يزداد زيادة مصطنعة أو زائفة، وأى خطأ أو صواب فى فهم الفقرة قد يؤثر فى الإجابة على نصفى الاختبار، ويؤدى ذلك إلى ارتفاع معامل الارتباط بين النصفين دون أن يدل فى واقع الأمر على ثبات مرتفع.

لقد بينا أننا نحصل على معامل الاتساق الداخلى بحساب معامل الارتباط بين درجات نصفى الاختبار إلا أننا يجب أن نشير إلى أن مثل هذا المعامل يعطينا ثباتا لنصف الاختبار فقط وليس للاختبار كله. وهذه المشكلة لاتوجد بالطبع فى الطريقتين السابقتين (إعادة الاختبار والصور المتكافئة).

ونحن نعلم أنه إذا تساوت الظروف الأخرى فإن الاختبار الأطول فى عدد أسئلته يكون أعلى ثباتا، لأننا لو حصلنا على عينة أكبر من السلوك نصل إلى مقاييس أكثر استقرارا، ويمكن تقدير الأثر الذى تحدثه إطالة الاختبار أو اختصاره على معامل الثبات باستخدام المعادلة التنبؤية الشهيرة التى اقترحها سبيرمان و براون Spear-man-Borwn وهى:

$$ر = \frac{ن\ ر_{1}}{1 + (ن - 1)\ ر_{1}}$$

حيث أن ر = معامل الثبات الذى نريد الحصول عليه

$ر_{1}$ = معامل الثبات الذى حصلنا عليه

ن = عدد مرات إطالة الاختبار أو اختصاره أى أنه إذا زاد عدد الأسئلة من ٢٥ إلى ١٠٠ سؤال فإن ن تكون ٤ إذا اختصر من ٦٠ إلى ٣٠ سؤال فإنها تصبح $\frac{1}{2}$.

ويمكن اختصار المعادلة حينما تطبق على معامل الثبات المقدر بطريقة التجزئة

النصفية لأنها تتضمن إطالة الاختبار مرتين، فتصبح على النحو الآتى:

$$ر = \frac{٢ ر_{١}}{١ + ر_{١}}$$

إلا أن معادلة سبيرمان – براون تفترض أن التغيرات فى درجات نصفى الاختبار متساوية تماما أى تفترض تكافؤ ثباتهما، وهو افتراض يصعب فى كثير من الأحوال توافره حتى ولو بدا نصفا الاختبار على درجة كبيرة من التكافؤ بالفعل، وقد اقترح رولون Rulon تبسيط معادلة سبيرمان – براون مع وضع تباين نصفى الاختبار فى الاعتبار، وصاغ لذلك المعادلة الآتية:

$$ر = ١ - \left[\frac{ع^{٢}ف}{ع^{٢}ك}\right]$$

حيث ع٢ف = تباين فروق درجات نصفى الاختبار.

ع٢ك = تباين الاختبار ككل.

وقد يستعاض عن ذلك باستخدام معادلة جتمان Guttman التى تستبعد هذا الافتراض وتصلح لحساب معامل الثبات مباشرة دون حاجة إلى حساب معامل الارتباط، ويمكن صياغة هذه المعادلة على النحو الآتى:

$$ر = ٢ [١ - \left(\frac{ع^{٢}_{١} + ع^{٢}_{٢}}{ع^{٢}ك}\right)]$$

وفى هذه المعادلة الرمز ع٢١، ع٢٢ = تباين النصف (١) والنصف (٢) فى الاختبار

ع٢ك = تباين الاختبار ككل.

٤ – طريقة كيودر – ريتشاردسون:

تصلح طريقة التجزئة النصفية لحساب معامل الاتساق الداخلى (الاتساق بين نصفى الاختبار)، أما بالنسبة لموضوع تجانس الأسئلة ذاتها والتى تؤلف فى مجموعها بنية الاختبار فإن بعض العلماء يرون أنه لايدخل فى نطاق مناقشة موضوع ثبات الاختبار وإنما هو من خصائص الاختبارات الجيدة بصفة عامة إلا أن هذا الرأى لايتفق مع التصور العام لمفهوم تباين الخطأ الذى عرضناه فيما سبق، فنحن ، دون

شك، فى حاجة إلى معرفة مدى استقرار استجابات المفحوص على مفردات الاختبار سؤالا بعد آخر وهذا ما يسمى «الاتساق بين أسئلة الاختبار» interitem consistency، وهو مصطلح آخر قد لا يعنى نفس ما يعنيه مصطلح معامل الاتساق الداخلى فى طريقة التجزئة النصفية، وحتى لايحدث الخلط بين الطريقتين نسمى مصدر التباين هنا بتجانس الاختبار homogeneity، وهو مصدر هام لأن عدم تجانس أسئلة الاختبار وتنوعها الشديد قد يؤدى إلى اختلاف أداء المفحوصين وتغيره فى المواقف المختلفة، فقد يوجد لدينا اختباران لهما معاملات ثبات مرتفعة حسبت اما بالصور المتكافئة أو بالتجزئة النصفية ولكنهما قد يختلفان فى المعاملات الخاصة بدرجة تجانس الأسئلة فيهما، وقد تظهر هذه المشكلة فى صورتى الاختبار أو نصفيه.. والواقع أن الفرق بين معامل الاتساق الداخلى (الذى يحسب بالتجزئة النصفية كما سبق أن أوضحنا) وبين معامل التجانس الذى نشير إليه هنا يمكن اعتباره دليلا على مدى التجانس أو التباين فى محتوى الاختبار.

وأشهر الطرق لحساب معامل التجانس طريقة كيودر- ريتشارد سون وهذه الطريقة تتشابه مع طريقة التجزئة النصفية فى أنها تستخدم إجراء اختباريا واحدا لصورة واحدة من الاختبار، ولكنها بدلا من أن تتطلب تجزئة الاختبار إلى نصفين متكافئين فإنها تعتمد على فحص أداء المفحوص على كل سؤال على حدة، وأشهر المعادلات الإحصائية التى تستخدم فى هذا الصدد هى:

$$ر = \left(\frac{ن}{ن-١}\right)\frac{ع^{٢} - مجـ ص خ}{ع^{٢}}$$

حيث أن ر = معامل ثبات الاختبار

ن = عدد الأسئلة التى تكون منها الاختبار.

ع٢ = تباين الاختبار ككل.

مجـ ص خ = يحسب باعداد نسبة الأفراد الذين أصابوا فى إجاباتهم على كل سؤال ونسبة الذين أخطأوا فيه ثم نضرب النسبتين فى بعضهما ونجمع حواصل الضرب بالنسبة لجميع الأسئلة.

ويؤكد كرونباك (٢٤) أن معامل التجانس كما يحسب بطريقة كيودر-

ريتشاردسون هو فى الواقع متوسط الاتساق الداخلى الناتجة عن تجزئات متعددة للاختبار، أما معامل الاتساق الداخلى فيعتمد على تجزئة تهدف إلى الحصول فقط على مجموعتين متساويتين أو متكافئتين من الأسئلة. والتى تعطى معاملا للثبات بالرغم من عدم تجانس الاختبار، ولعل من الأمثلة المتطرفة على ذلك أن يكون كل سؤال زوجى وفردى فى الاختبار يقيس عملية نفسية مختلفة عن أى سؤال فردى وزوجى آخر فى الاختبار.

ومن الواضح أن معادلة كيودر- ريتشاردسون تصلح للاختبارات التى تتم الإجابة على أسئلتها بالثنائية، من نوع الكل أو لا شىء مثل صواب - خطأ، نعم - لا، موافق - معارض وغيرها، ولهذا تقترح آمال صادق وفؤاد أبو حطب (١) أن يحل الرمزان (أ) ، (ب) محل (ص)، (خ) فى المعادلة الأصلية لتصبح أوسع فى الاستخدام وتمتد إلى نطاق اختبارات الأداء المميز.

٥ - معامل ألفا لكرونباك:

كثير من الاختبارات النفسية لايقوم على نظام الكل أو لا شىء المشار إليه عند الحديث عن طريقة كيودر- ريتشاردسون، ففى التقرير الذاتى قد يحصل المفحوص على درجة مختلفة تبعا لدرجة إجابته على سؤال يطلب تقدير مدى حدوث سلوك معين بأنه يحدث «عادة» أو «أحيانا» أو «قليلا» أو «نادرا»، أو عند الإجابة على سؤال فى مقياس للاتجاهات بقوله أزاء عبارة معينة بأنه «موافق جداً» أو «موافق» أو «لا رأى له» أو «معارض» أو «معارض جداً»، وبالطبع فإن هذا النوع من الاختبارات قد يكون أكثر حاجة إلى تحديد معامل التجانس فيه أكثر من اختبارات الأداء الأقصى التى هى فى العادة من نوع الكل أو لا شىء، ولهذا اقترح كرونباك منذ عام ١٩٥١ معادلة تطورت بعد ذلك على يد كايزر وميشيل عام ١٩٧٥ أطلق عليها معامل ألفا Alpha Coefficient، وفيها يحل مجموع تباينات درجات الأسئلة محل مجـ ص خ (أو مج أ ب إذا شئت) وصيغت المعادلة كما يلى:

$$ر = \left(\frac{ن}{ن - ١}\right) \frac{ع^{٢} - مجـ (ع^{٢}\ ١، ٢، ن)}{ع^{٢}}$$

حيث الرمز مجـ ع٢ ١، ٢، ٣، ...، ن = مجموع تباينات الأسئلة من السؤال الأول فى الاختبار إلى السؤال الأخير.

٦ - ثبات الفاحصين والمصححين:

بعد أن نستبعد الأنواع الثلاثة السابقة من تباين الخطأ من التباين الكلى فإن التباين المتبقى قد يرجع إلى الفروق الفردية الحقيقية موضوع الاهتمام، وبعضها الآخر إلى عوامل أخرى غير مرتبطة بالاختبار ولكن يمكن ضبطها تجريبيا، فمثلا لايسجل الباحثون أخطاء القياس الناتجة عن تطبيق الاختبار فى ظروف مشتتة للانتباه، أو فى زمن أطول أو أقصر مما يحدده كتيب الاختبار، أو إعطاء تعليمات مختلفة عن التعليمات المحددة، لأن هذه الأخطاء جميعا يمكن استبعادها والتغلب عليها فى الموقف الاختبارى نفسه إذا أتقن الفاحص فن إعطاء الاختبارات الذى عرضناه فى الفصل الثالث. ولذلك فنحن لانجد فى العادة معاملات ثبات خاصة تشير إلى «التباين الخاص بتشتت الانتباه» أو «تباين اختلاف الزمن» أو «تباين اختلاف التعليمات» إلخ. فمعظم الاختبارات تحدد إجراءات مقننة للتطبيق (أشرنا إليها فى الفصل الثالث) والتصحيح (سنشير إليها فى الفصل السادس) بحيث نفترض مع وجود هذه الإجراءات والالتزام بها أن «ثبات الفاحص» و«ثبات المصحح» يصل إلى درجة كبيرة من الكفاءة تجعل الاختبار مفيدا فى الأغراض التطبيقية والعملية وبالتالى لانكون فى حاجة إلى تقدير مثل هذا النوع من الثبات الذى يرتبط فى وجوهره بموضوعية الملاحظة أو التصحيح، ويصدق هذا القول على وجه الخصوص بالنسبة للاختبارات الموضوعية الجماعية التى نطبقها على جماعات كبيرة والتى تصححها الحاسبات الالكترونية، ففى هذه الحالات جميعا لا يتطلب الأمر إلا التأكد من أن الفاحص يتتبع إجراءات التطبيق والتصحيح تتبعا دقيقا وتكون المشكلة فى جوهرها الضبط المحكم لظروف إجراء الاختبار.

إلا أننا نلاحظ أنه فى الاختبارات الفردية مثل ستانفورد - بينيه يلعب الممتحن دوراً أكثر تعقيداً، لأن إجراءات الاختبارات فى هذه الأحوال ليست مقننة تقنينا واحدا بل يعتمد معظمها على كفاءة تدريب الفاحص، وعلى نجاحه فى تكوين علاقة بينه وبين المفحوص، وعلى دوافع المفحوص، وقد يتطلب الأمر أن يقوم الفاحص بتقويم أداء المفحوص خلال عملية تطبيق الاختبار لأن مثل هذا الأداء يحدد كيف يستمر الفاحص فى موقف الاختبار، ولذلك فقد يختلف الفاحصون فى النتائج التى يحصلون عليها من نفس المفحوصين حتى ولو توافرت فى هؤلاء الفاحصين أفضل الشروط،

وهذا الاختلاف قد يكون من نوع تباين الخطأ الذى يمكن أن ترجعه إلى الفروق الفردية بين الفاحصين ويؤثر بالطبع على التباين الكلى لدرجات المفحوصين، ولذلك فإذا كانت الاختبارات التى نستخدمها من النوع الذى تلعب فيه فردية الفاحص دوراً مهما، أو إذا كانت من النوع الذى يتأثر بالعوامل الذاتية وطرق الملاحظة المختلفة، نكون فى حاجة ماسة إلى تقدير «ثبات الفاحص» وخاصة إذا كان من الضرورى معرفة العلاقة بين تقديرات أكثر من فاحص لنفس المفحوص أو المفحوصين.

وبالمثل نجد أن بعض الاختبارات تثير مشكلة «ثبات المصحح» وخاصة اختبارات المقال والاختبارات الإسقاطية بوجه عام، وهى الأدوات التى تسمح بمقدار كبير من التفسير الذاتى من جانب المصحح وهو الفاحص أيضاً فى كثير من الأحوال.

ومن الواضح أن معامل ثبات المصحح يمكن الحصول عليه بحساب معامل الارتباط بين الدرجات التى يعطيها مصححان أو أكثر لنفس الأفراد فى نفس الاختبارات، ومعنى ذلك أن كل مفحوص يحصل على درجتين أو أكثر من تصحيح اختبار واحد. وهذه الطريقة لا تتضمن بالطبع نواحى القصور الموجودة فى طريقة إعادة الاختبار، أما معامل ثبات الفاحص فهو أكثر صعوبة لأنه بطبيعته يتضمن إعادة اختبار من جانب أكثر من فاحص، وبالتالى فإن قيام أكثر من فاحص بإعادة الاختبار يعنى الوقوع فى جوانب القصور التى سبق الحديث عنها فى طريقة إعادة الاختبار التى يقوم بها نفس الفاحصين، بالإضافة إلى أن معامل الارتباط الناتج يعكس عوامل الاستقرار الوقتى وثبات المصحح معا، ويمكن اللجوء إلى نمط مختلف من التخطيط التجريبى يعتمد على تحليل التباين لعزل تباين الخطأ الذى يمكن إرجاعه إلى الفاحصين، وعلى كل فيمكن أن نعتبر كلا من معامل ثبات المصحح ومعامل ثبات الفاحصين. من نوع واحد ونطلق عليه اسم معامل الموضوعية Coefficient of objectivty.

ويمكن أن نلخص أنواع معاملات الثبات وطرق تقديرها فى الجدول رقم (٩).

جدول رقم (٩) أنواع تباين الخطأ وعلاقته بطرق حساب الثبات (عن ١٣ : ١١٨).

المعامل	طريقة تقديره	تباين الخطأ
معامل الاستقرار	إعادة الاختبار بنفس صورته على نفس المفحوصين فى مناسبة مختلفة	عينة الوقت
معامل التكافؤ	تطبيق الاختبار مع صورة مكافئة له فى نفس الوقت	عينة المحتوى
معامل الاستقرار والتكافؤ معا	تطبيق الاختبار ثم تطبيق صورة مكافئة له فى مناسبة مختلفة	عينة الوقت ، عينة المحتوى
معامل الاتساق الداخلى	التجزئة النصفية	عينة المحتوى، تجانس المحتوى
معامل التجانس	تحليل المفردات بطريقة كيودر- ريتشاردسون أو بطريقة ألفا	عينة المحتوى، تجانس المحتوى
معامل الموضوعية	تكرار المصحح أو الفاحص	ذاتية الفاحص ، عينة الوقت

العوامل المؤثرة فى ثبات الاختبار:

١ - درجة اعتماد الاختبار على الزمن والسرعة.

تمثل الاختبارات ذات الزمن المحدد مشكلة من نوع خاص فى موضوع الثبات، وخاصة إذا كانت الفروق الفردية فى درجات هذه الاختبارات تعتمد على سرعة الأداء. ولذلك لايمكن استخدام طريقتى التجزئة النصفية أو كيودر- ريتشارد سون لحساب ثبات هذه الاختبارات لأن هاتين الطريقتين تعتمدان على الاطراد فى ارتكاب الأخطاء، فمثلا إذا كان الاختبار يتكون من ٥٠ سؤالا من النوع الذى يعتمد على السرعة اعتمادا كليا (أى على عدد الأسئلة التى يحاولها المفحوص فى جلسة الاختبار) كاختبار النقر أو اختبار العمليات العددية لطلاب الثانوية العامة مثلا فإننا نجد أن المفحوص الذى يحصل على الدرجة ٣٦ مثلا تكون درجته فى كل من النصف الفردى والنصف الزوجى من الاختبار ١٨ ، بعبارة أخرى تتساوى تقريبا درجته فى نصفى الاختبار إلا إذا تأثرت هذه الدرجة بالأخطاء العارضة التى ترجع إلى إهماله لعدد قليل من هذه الأسئلة، وفيما عدا ذلك فإن الارتباط بين الدرجات الفردية والدرجات الزوجية يكون أقرب إلى الارتباط التام، وهو معامل زائف لايدل بحال من الأحوال على ثبات مرتفع الاختبار.

ولذلك فأفضل طرق حساب الثبات لهذه الاختبارات هى طريقتا إعادة الاختبار

واستخدام الصور المتكافئة، ويمكن اللجوء إلى طريقة التجزئة النصفية بشرط أن تتم التجزئة على أساس السرعة فى الأداء لا على أساس عدد الأخطاء، بعبارة أخرى يجب أن تعتمد درجات نصفى الاختبار على مجموعات من الأسئلة ذات زمن محدد منفصل بعضها عن بعض، ويمكن تحقيق ذلك باستخدام جزئين متكافئين من الاختبار يعطى كل جزء منهما فى نصف الزمن الكلى للاختبار، وهذه الطريقة أقرب إلى طريقة تطبيق الصور المتكافئة من الاختبار فى الجلسة نفسها أو بعدها مباشرة، إلا أن كل صورة هى نصف اختبار فى حين أن الدرجة التى يحصل عليها المفحوص هى درجته فى الاختبار الكلى (أى فى نصفيه)، ولذلك لابد من استخدام معادلة سبيرمان - براون أو غيرها من المعادلات المناسبة لإيجاد ثبات الاختبار ككل.

وإذا استحال تطبيق نصفين للاختبار مستقلين فيمكن تقسيم الزمن الكلى إلى أكثر من نصفين ثم إعطاء درجة لكل قسم، ويمكن أن يتم ذلك بإنن نطلب من المفحوص أن يضع علامة معينة أمام السؤال الذى يكون قد وقف عنده عندما يعطى الممتحن إشارة معينة متفق عليها مقدما، ويمكن أن نجمع عدد المفردات التى أجاب عليها المفحوص إجابة صحيحة فى القسمين الأول والرابع مثلا ونعتبرها مجموعة مستقلة ثم نجمع إجابات القسمين الثانى والثالث فى مجموعة أخرى بحيث يمثل كل من هاتين المجموعتين نصف اختبار، ومثل هذا النوع من التقسيم الرباعى يحقق نوعا من التوازن بين الآثار التراكمية للممارسة والتعب وغيرها من العوامل، وتصلح هذه الطريقة فى الاختبارات الأخرى التى تعتمد على السرعة وخاصة حينما لاتكون الأسئلة مرتبة حسب مستويات الصعوبة.

والسؤال هو: متى يمكن أن نعتبر الاختبار من النوع الموقوت وبالتالى يجب أن نأخذ فى اعتبارنا ما سبق مناقشته؟ من الواضح أن مجرد استخدام زمن محدد للاختبار لايعنى أنه اختبار سرعة لأنه إذا انتهى جميع المفحوصين من الإجابة فى الزمن الذى يحدده الاختبار فإن السرعة لا تلعب دوراً فى تحديد الدرجة، وحتى إذا أتم جميع المفحوصين ٤٠ سؤالا من ٥٠ فإن الفروق الفردية التى ترجع إلى السرعة تصبح لاوجود لها رغم أنه لايوجد مفحوص واحد أجاب على جميع أسئلة الاختبار، ولذلك لايمكن أن نعتبر مجرد محاولة الاختبار كله أو عدم محاولته دليلا على السرعة فى مقابل القوة، وإنما لابد من اللجوء إلى وسائل أكثر دقة لتحديد المدى الذى يمكن

أن نرجع فيه الفروق الفردية فى الأداء الاختبارى إلى السرعة ويتم ذلك بتحديد النسبة التى يمكن أن نرجعها إلى التباين فى السرعة من التباين الكلى لدرجات الاختبار، ويمكن تحديد هذه النسبة بسهولة بإيجاد تباين درجات الأسئلة التى أجاب عليها جميع المفحوصين وقسمة هذا المقدار على تباين الدرجات الكلية للاختبار ويمكن صياغة ذلك فى صورة المعادلة الآتية:

$$\text{(ع}^2\text{س) تباين السرعة} = \frac{\text{تباين الدرجات التى تمت الإجابة عليها (ع}^2\text{ص)}}{\text{التباين الكلى لدرجات الاختبار (ع}^2\text{ك)}}$$

فإذا حدث وأجاب جميع المفحوصين على ٤٠ سؤالا يصبح بسط المعادلة صفرا حيث لاتوجد فروق فردية (أو تباين) فى عدد الأسئلة التى تمت الإجابة عليها (ع^{2}ص = صفر)، وبالتالى يصبح ناتج المعادلة مساويا للصفر ولايتحقق ذلك بالطبع فى اختبارات القوة والتى لا تتأثر مطلقا بالسرعة، ومن ناحية أخرى إذا كان التباين الكلى للاختبار (ع^{2}ك) يرجع إلى فروق فردية فى السرعة فقط ولايتأثر مطلقا بالقوة فإن تباين الدرجات التى قد تكون تمت الإجابة عليها (ع^{2}ص) يصبح مساويا للتباين الكلى ويكون ناتج المعادلة مساويا للواحد الصحيح أى يكون الاختبار اختبار سرعة بحت.

وهناك طرق إحصائية أخرى عديدة وأكثر دقة لتحديد هذه النسبة ولكن مناقشتها تفصيلا يخرج عن نطاق هذا الكتاب.

ويمكن أن نضرب مثلا بأثر السرعة على معاملات التناسق الداخلى من بحث قامت به أنستازى ودربك (فى ١٣ : ٥٢ – ١٢٣) على اختبارات القدرات العقلية الأولية لأعمار تتراوح بين ١١ سنة و١٧ سنة، وفى هذا البحث حسبت معاملات الثبات بطريقة التجزئة النصفية أولا ثم عن طريق حساب معاملات ارتباط الدرجات فى نصفين منفصلين تبعا للسرعة والزمن وتتلخص النتائج فى الجدول رقم (١٠).

جدول رقم (١٠) معاملات الثبات لبعض اختبارات بطارية القدرات العقلية الأولية

طريقة حساب معامل الثبات	معانى الكلمات	الاستدلال	المكان	العدد
التجزئية النصفية فى محاولة واحدة	٠,٩٤	٠,٩٦	٠,٩٠	٠,٩٢
استخدام نصفين مستقلين من حيث الزمن	٠,٩٠	٠,٨٧	٠,٧٥	٠,٨٣

ومن هذا الجدول يتضح أن اختبار معانى الكلمات هو فى أساسه اختبار قوة، بينما يعتمد اختبار الاستدلال على السرعة إلى حد ما، أما اختبار المكان فيعتمد اعتمادا كبيرا على السرعة، كما يتضح أيضاً أننا عندما نحسب معاملات الثبات بطريقة جيدة فإن هذه المعاملات تنخفض فى اختبار المكان من٠,٩٠ إلى ٠,٧٥، وكذلك الشأن فى اختبارى الاستدلال والعدد فإن معلملات ثباتها تنخفض نسبيا، أما اختبار معانى الكلمات الذى لا يعتمد على السرعة فإن معامل ثباته لم يختلف اختلافا له دلالة عندما حسب بكل من الطريقتين آنفتى الذكر.

أما إذا حسب معامل ثبات الاختبارات الموقوتة بطريقة التجزئة النصفية، ثم صحح أثر التجزئة بمعادلة سبيرمان ـ براون فإن جلكسون Gulkson يقترح تصحيحا إضافيا بالمعادلة الآتية:

$$ر = ر' - \left(\frac{م\ ت}{ع^{1}\ جـ}\right)$$

حيث الرمز ر = معامل الثبات بعد تصحيح أثر السرعة

ر' = معامل الثبات بعد تصحيح سبيرمان ـ براون

م ت = متوسط الأسئلة المتروكة

ع^{2} حـ = تباين الخطأ (تباين الإجابات الخاطئة والإجابات المحذوفة).

٢ – مدى الفروق الفردية فى العينة:

تمثل طبيعة الفروق الفردية فى عينة الأفراد التى تطبق عليها الاختبارات

بغرض تقدير معاملات الثبات عاملا مهما مؤثرا فى هذه المعاملات، ويجب أن نؤكد أن معامل الارتباط عموما يتأثر بمدى الفروق الفردية فى الأفراد، فإذا انعدمت هذه الفروق بالنسبة لخاصية من الخصائص فإن معامل ارتباط هذه الخاصية بغيرها يصبح صفرا، ويكون من المستحيل علينا أن نتنبأ بمركز الفرد فى أى خاصية أخرى من معرفة درجته فى الاختبار الذى وضعناه، وإذا أخذنا مثالا على ذلك أقل تطرفا وليكن معامل الارتباط بين درجات اختبار للفهم اللغوى واختبار للاستدلال الحسابى طبقا على عينة متجانسة تجانسا شديدا (مثلا طلاب إحدى كليات الجامعة) فإن معامل الارتباط فى هذه الحالة قد يكون منخفضا انخفاضا واضحا دون أن يدل ذلك على ضعف العلاقة بين الاختبارين، أما إذا أعطينا الاختبارين لعينة غير متجانسة فإن معامل الارتباط بينهما قد يكون أعلى مده فى الحالة الأولى.

ويوضح لنا ذلك أثر تحديد مدى الفروق الفردية على معامل الارتباط بوجه عام، ويطلق على هذه الظاهرة فى علم الإحصاء اسم المدى المحدود للعينة restricted range ، ويعتمد ذلك فى جوهره على مدى التباين فى العينة.

وهذا القول يصدق فى عمومه على معاملات الثبات - باعتبارها فى الأصل معاملات ارتباط - وتوجد معادلات كثيرة لتقدير معامل الثبات المتوقع عندما يزيد أو يقل الانحراف المعيارى فى الجماعة.

٣ - خصائص العينة:

يختلف معامل الثبات بين الجماعات التى يتفاوت متوسط مستويات القدرة فيها، وهذه الفروق لايمكن التنبؤ بها أو تقديرها على جماعات تختلف فى العمر أو فى مستوى القدرة، ومثل هذه الفروق قد تنشأ من أن المستويات المختلفة لصعوبة الاختبار تقيس فى الواقع قدرات مختلفة نوعا ما بعضها عن بعض، أو قد تنشأ من الخصائص الإحصائية لنوع الدرجة الذى نستخدمه فى تصحيح الاختبار كما هو الحال فى درجة نسبة الذكاء فى اختبار ستانفورد - بينيه مثلا، حيث نجد أن معاملات الثبات قد تكون منخفضة نسبيا فى الجماعات الأصغر عمرا أو الأقل قدرة لأن هذه الدرجات تتأثر فى هذه الجماعات بالتخمين، ولذلك لايجب تطبيق الاختبار فى هذه المستويات.

ومن الواضح أن كل معامل ثبات يجب أن يصحبه وصف دقيق لنوع المجموعة

التى حسب من درجات أفرادها، ويجب أن ننبه على وجه الخصوص إلى تباين القدرة ومستوياتها فى العينة، ولايكون معامل الثبات الذى يتضمنه دليل الاختبار صالحا إلا للجماعات المشابهة لعينة التقنين، ولذلك يحسن أن نقسم عينه التقنين هذه إلى مجموعات فرعية أكثر تجانسا بالنسبة لعوامل الجنس والعمر والصف الدراسى (أو المرحلة التعليمية) والمهنة وغيرها، ثم حساب معاملات ثبات منفصلة لكل من هذه المجموعات، وفى هذه الحالة يمكن الاستفادة من معاملات الثبات التى حسبت من عينات التقنين حين نستخدم الاختبار فى الأغراض التطبيقية والعملية.

٤ - عدد أسئلة الاختبار:

يزداد معامل ثبات الاختبار مع زيادة عدد الأسئلة التى تكونه، ولعل هذا العامل لايحتاج إلى مزيد من التفصيل بعد مناقشتنا المستفيضة السابقة لتصحيح أثر تجزئة الاختبار فى طريقة التجزئة النصفية، حيث يفترض أن معامل ثبات نصف الاختبار أقل من معامل ثبات الاختبار الكلى.

٥ - تباين الخطأ:

يؤثر فى ثبات الاختبار أى مصدر لتباين الخطأ يكون فعالا فى الموقف الاختبارى، ومفهوم الثبات يعتمد فى جوهره كما بينا آنفا على مدى تحرر الاختبار من هذه المصادر، وعلى الباحث النفسى أن يجرب كل سبيل ممكن للتأكد من ذلك ولعل طرق حساب الثبات التى عرضناها تهيىء له فرصاً متعددة متنوعة للوصول إلى هذا الهدف.

الخطأ المعيارى للقياس:

يمكن التعبير عن مفهوم الاختبار فى ضوء الخطأ المعيارى للقياس (ع ق)، أو الخطأ المعيارى للدرجة فى الاختبار، ويفيد هذا المفهوم فائدة كبرى فى تفسير درجات المفحوصين كأفراد، ولذلك تفوق أهميته فى بعض الحالات مفهوم الثبات نفسه، ويمكن حساب الخطأ المعيارى للقياس من معامل الثبات مباشرة بالمعادلة الآتية:

$$ع_{ق} = ع_{ت} \sqrt{١ - ر}$$

حيث ع ق = الخطأ المعيارى للقياس

ع ت = الانحراف المعيارى لدرجات الاختبار،

ر = معامل الثبات

بشرط أن يحسب ع ت ، ر لنفس العينة.

ولتوضيح ذلك نعطى مثالا : نفرض أن الانحراف المعيارى لأحد اختبارات القدرة الموسيقية ١٠ ومعامل ثباته ٠,٨٤ فإننا نحصل على الخطأ المعيارى للقياس مباشرة بالمعادلة السابقة كما يلى:

ع ق = ١٠ $\sqrt{١ - ٠,٨٤}$ = ٤

ويمثل هذا المقدار الدرجة التى تتذبذب فيها درجات المفحوص إذا افترضنا أنه تعرض للقياس بنفس الاختبار عددا كبيرا من المرات وأن درجاته تتوزع توزيعا اعتداليا، أى أنه يعبر عن الانحراف المعيارى لدرجات المفحوص الواحد فى نفس الاختبار، لو أعيد تطبيق هذا الاختبار عليه عشرات المرات.

لنفرض أن أحد المفحوصين حصل على الدرجة ٤٥ فى هذا الاختبار ونريد أن نحدد الحدود الدنيا والعليا التى تتوقع لدرجاته ألا تتجاوزها لو أعيد عليه القياس (حدود الثبات)، فى هذه الحالة يقوم الباحث بما يلى:

١ – اختيار إحدى نقاط المنحنى الاعتدالى ولتكن النقطة التى تزيد على المتوسط (م) بما يعادل + ٢,٥٨ ع، وهى النقطة التى تحددها نظرية المنحنى الاعتدالى بأن عندها يمكن أن يتم استيعاب ٩٩ ٪ من درجات هذا المفحوص لو أعيد تطبيق الاختبار عليه عشرات المرات، أما النسبة المئوية الباقية (وهى ١ ٪) فتمثل احتمال حدوث خطأ القياس وعدم حصول المفحوص على الدرجة، أى أن حدوث احتمال حدوث الخطأ فى هذه الحالة ± ١ ٪.

٢ – حساب النقطة التى تقابل (٢,٥٨ ع) فى درجات المفحوص بافتراض أن الخطأ المعيارى (ع ق) هو فى جوهره انحراف معيارى (ع) وفى مثالنا الحالى تكون هذه النقطة = ± ٢,٥٨ × ٤ = ± ١١,٣٢

والتى قد تكون موجبة أو سالبة حسب موضع الانحراف المعيارى من المتوسط.

ومن هذا المثال يمكن أن نستنتج أن درجة هذا المفحوص عند تطبيق الاختبار

عليه مرة واحدة سوف تمتد بين (٤٥ – ١١) و (٤٥+١١) أى بين ٣٤، ٥٦ بدرجة من اليقين يصل إلى ٩٩٪، فإذا طبق عليه هذا الاختبار أو أى اختبارات مكافئة مائة مرة فإن درجته التى يمكن أن تقع خارج هذا النطاق لن تحدث إلا مرة واحدة، ويمثل هذا النطاق حدود الدرجة الحقيقية true score للفرد.

ومن الواضح أن الخطأ المعيارى للقياس يصف الأداء الفردى ويفسره بينما يعبر معامل الثبات عن الأداء الجماعى، ولهذا لاتصلح الأخطاء المعيارية للمقارنة بين الاختبارات فى درجة الدقة، وفى هذه الحالة لا تصلح إلا معاملات الثبات.

ولعل من أهم فوائد الخطأ المعيارى للقياس - إلى جانب تفسيره لأداء الفرد - أنه يجعل الفاحصين لايبالغون فى تقدير الفروق الضئيلة بين الدرجات وهذا الحذر يفيد الفاحص سواء عند المقارنة بين درجات مفحوصين مختلفين فى الاختبار الواحد أوالمقارنة بين درجات نفس المفحوص فى اختبارات مختلفة، وكذلك يجب فهم الدرجات التى يحصل عليها المفحوصون بعد المعالجات التجريبية فى ضوء الخطأ المعيارى للقياس، إلا أننا فى هذهالحالة نحتاج إلى حساب الخطأ المعيارى للفروق بين المقاييس باستخدام المعادلة الآتية:

$$ع_{ف} = \sqrt{ع_{ق_1}^2 + ع_{ق_2}^2}$$

حيث ع ف = الخطأ المعيارى للمقياس الأول والمقياس الثانى على التوالى.

ويمكن إعادة صياغة المعادلة السابقة باستخدام معاملات الثبات كما يلى:

$$ع_{ف} = ع\sqrt{٢ - ر_1 - ر_2}$$

مع ملاحظة أن الرمز ع يستخدم للاختبارين معا لأن درجاتهما يجب أن تحول إلى نفس المقياس قبل المقارنة بينهما (أى يتم تحويل الدرجات الخام إلى درجات معيارية أو درجات معيارية معدلة على النحو الذى سنبينه فى الفصل السابع). ويوضح ذلك المثال الآتى:

نفرض أننا طبقنا اختبارين معامل ثباتهما ٠,٩٧، ٠,٩٢ على التوالى، وأن درجاتهما الأصلية الخام حولت إلى نسب ذكاء انحرافية (درجات معيارية معدلة) متوسطها ١٠٠ وانحرافها المعيارى ١٥، يمكن أن نحسب الخطأ المعيارى للفرق بين

درجتى هذين الاختيارين بالتعويض فى المعادلة السابقة كما يلى:

ع ف = ١٥ $\sqrt{}$ ٢ – ٩٧, – ٩٣, = ٤,٧٤

ويمكن أن نستخدم هذه القيمة بنفس المعنى الذى استخدم فيه الخطأ المعيارى للمقياس الواحد (Anastasi & urbna 1997 p. 111).

الفصل الخامس

صدق الاختبارات

معنى صدق الاختبار

يتعلق موضوع صدق الاختبارات Validity بما يقيسه الاختبار وإلى أى حد ينجح فى قياسه ويتصل هذا بمدى وصولنا إلى تنبؤ دقيق أو استنتاج صحيح من الدرجة التى يحصل عليها المفحوص فى الاختبار، وهذا الموضوع لايقتصر على عملية القياس وحدها وإنما قد يمتد إلى المنهج التجريبى بصفة عامة والذى يتطلب اختبار فروض علمية حول طبيعة الاختبار وما يقيسه.

وإذا كان الصدق مسألة على درجة كبيرة من الأهمية فى تقرير صلاحية أو فائدة أى فرض علمى أو فى تحديد معنى مفهوم معين فإنه فى الواقع هو الذى يحدد قيمة الاختبار ومغزاه، ويجب أن نقرر منذ البداية أن عنوان الاختبار لايدل بالضرورة على ما يقيسه الاختبار لأنه عادة ما يكون مجرد كلمة أو عبارة قصيرة تهدف إلى التعرف على الاختبار. ولايمكن أن نستدل على العملية النفسية التى يقيسها الاختبار إلا بفحص مجموعة من المعلومات والبيانات التى تتوفر لنا بالطرق الصحيحة وتكون مستقلة نسبيا عن الاختبار.

طرق تحديد الصدق:

يمكن القول بوجه عام أن جميع طرق تحديد صدق الاختبارات تبحث عن العلاقة التى تربط أداء الأفراد فى الاختبار وبعض الحقائق والمعلومات التى تتصل بالعملية النفسية موضوع الاختبار ، والتى (أى هذه المعلومات والحقائق) يمكن ملاحظتها مستقلة، وهناك كثير من الطرق التى تستخدم فى هذا الصدد ولها مسميات كثيرة، إلا أننا سوف نتحدث عن الأنواع الأربعة من الصدق التى وصفها كتيب التوصيات الفنية للجمعية الأمريكية لعلم النفس (١٦، ١٧) والتى عالجتها مؤلفات أنستازى (١٣) وكرونباك (٢٣) بالتفصيل.

١ - صدق المحتوى:

يقصد بصدق المحتوى Content Validity أن يقوم الباحث النفسى بفحص

مضمون الاختبار فحصاً دقيقا منتظما لتحديد ما إذا كان يشتمل على عينة ممثلة لميدان السلوك الذى يقيسه كما هو الحال فى الاختبارات المرجعة إلى المعيار أو تحتوى على معظم إن لم يكن جميع جوانب هذا السلوك، كما هو الحال فى الاختبارات المرجعة إلى المحك، ومثل هذا النوع من الصدق نستخدمه فى العادة فى تقويم الاختبارات التحصيلية التى نعدها لقياس مقدار ما تمكن المتعلمون من اتقانه من جوانب مهارة معينة أو ما حصلوه من مقرر دراسى أو برنامج مهنى، وقد يبدو أن مجرد فحص مضمون الاختبار يكفى للتدليل على صدقه وصلاحيته للغرض الذى وضع من أجله، فاختبار معين فى الضرب أو التاريخ قد يبدو صادقا بالتعريف إذا اشتمل على أسئلة تدور حول عملية الضرب أو التاريخ، إلا أن ما نحب أن ننبه إليه أن هذا قد لا يكون صحيحاً، فليس مجرد تأمل محتوى الاختبار يدل على صدقه.

ولكن الأمر ليس على هذا القدر من السهولة لأن من أهم المشكلات التى نواجهها هنا هى مشكلة عينة المحتوى أو أصله الإحصائى العام، فنطاق المحتوى الذى نختبره يجب أن يتم تحليله تحليلا دقيقاً منتظماً حتى نتأكد من أن جميع الجوانب الأساسية تغطيها أسئلة الاختبار تغطية ملائمة وبالنسب الصحيحة حتى لا يكون لدينا عدد كبير من الأسئلة فى جانب معين على حساب الجوانب الأخرى، ويتطلب ذلك بادئ ذى بدء، وصفا تفصيلياً دقيقا لمدى المهام والمثيرات والمواقف والمواد التى يستوعبها نطاق المحتوى ، كما يتطلب تفصيل أنواع الاستجابات التى يتوقع لها أن تصدر عن المفحوص ويجب أن يهتم بها الفاحص أو الملاحظ أو المصحح، والتى قد تدل على عمليات نفسيه معينة، وكذلك توصيف الأهداف التعليمية أو التدريبية أو العلاجية بشرط أن تصاغ فى عبارات إجرائية سلوكية تقبل الملاحظة والقياس، وبهذا يكون صدق المحتوى أكثر اعتمادا على مدى الارتباط بين استجابات المفحوص على أسئلة الاختبار وميدان السلوك الذى نختبره، ففى قياس القدرة على القراءة مثلا قد يلجأ الباحث - بالنسبة لأهداف معينة - إلى بناء اختباره فى حدود المادة القرائية التى تقدمها المدرسة، وقد يلجأ - بالنسبة لأهداف أخرى - إلى أن يتناول الاختبار نطاقا أوسع من المواد المقروءة مثل الروايات والصحف والمجلات، ولهذا لابد لمؤلف الاختبار من أن يحدد فى دليل اختباره وبدقة عناصر التحليل الثلاثة التى أشرنا إليها.

(المواد والاستجابات والأهداف) حتى يمكن لأى فاحص أن يحكم على مدى ملاءمة الاختبار لأهدافه فى التقويم النفسى.

وصدق المحتوى يفيد كما سبق أن قلنا فى اختبارات التحصيل، فهذه الاختبارات يكون لمحتواها - فى العادة - أصل إحصائى يمكن أن يعبر عنه كله فى الاختبار، أو تنتقى منه عينة ممثلة، وخاصة فى الاختبارات التحصيلية التى تستخدم فى تقويم مدى الوصول إلى الأهداف التعليمية من المستويات المنخفضة، إلا أن صدق المحتوى لايكون كافيا لاختبارات الاستعدادات ومقاييس الشخصية بل ولاختبارات التحصيل التى تقيس العمليات المعرفية العليا، بل قد يكون مضللا فى كثير من الحالات، فعلى الرغم من أن هذا النوع الأخير من الاختبارات لابد أن يتوفر لها محكا الملاءمة والفاعلية بالنسبة لمحتواها وخاصة فى المراحل الأولى لإعدادها إلا أن هذا وحده لايكفى، بل لابد من أن نلجأ بعد ذلك إلى وسائل التحقق التجريبى المضبوط، التى سوف نتحدث عنها فيما بعد، لتقدير صدقها، ويتضح ذلك إذا علمنا أن اختبارات الاستعدادات والعمليات المعرفية العليا والشخصية، على عكس الاختبارات التحصيلية، لا تتشابه تشابها داخليا كبيرا مع ميدان السلوك الذى تعد عينة منه، وبذلك لايكون لمحتواها أى دور فى التحقق من الفروض التى يصوغها مؤلف الاختبار عند انتقاء محتواه، لأن هذه الفروض تتطلب فى هذه الحالة تحقيقا تجريبيا منظما حتى يمكن التأكد من صدق الاختبار، وبعبارة أخرى نستخدم الإجراءات التجريبية المنظمة للتحقق من درجة التشابه بين عينة السلوك الذى يتضمنها اختبار الاستعداد والسلوك الذى تنتمى إليه.

ومن ناحية أخرى فإن اختبارات الاستعداد والشخصية على عكس الاختبارات التحصيلية لا تعتمد على مقرر دراسى أو على برنامج تدريس أو على مجموعة موحدة من الخبرات السابقة يمكن منها صياغة محتوى محدد للاختبار، وإنما يتفاوت الأفراد كثيراً فى طرق عملهم وفى العمليات النفسية التى يستخدمونها عند لإجابة على نفس الأسئلة، وعلى ذلك فإن الاختبار الواحد قد يقيس عمليات سلوكية مختلفة فى الأفراد المختلفين، وفى مثل هذه الأحوال يستحيل علينا تحديد هذه العمليات التى يقيسها الاختبار من مجرد فحص محتواه، فقد يحل طالب الجامعة مشكلة معينة مستخدما التعبير اللغوى (أو الرمزى عموما ومنه الرياضى)، بينما يحلل الميكانيكى

نفس المشكلة ويصل إلى نفس الحل مستخدما التصور البصرى المكانى. كما أن اختبارا معينا للاستدلال الحسابى لطلاب المرحلة الإعدادية قد يقيس مجرد الفروق الفردية فى سرعة العمليات العددية عند طلاب الجامعة.

ويجب ألا نخلط بين صدق المحتوى والصدق الظاهرى face validity فالصدق الظاهرى ليس صدقا بالمعنى العلمى للكلمة لأنه يدل على ما يبدو أن الاختبار يقيسه من الظاهر لا على ما يقيسه بالفعل، كما أنه يعنى ما إذا كان الاختبار يبدو صادقا فى نظر كل من المفحوين والفاحصين، والواقع أن موضوع الصدق الظاهرى يدخل من باب العلاقة الطيبة التى يكونها الفاحص مع المفحوص، وهو من الأمور المرغوب فيها دون شك، رغم ما يسببه من خلط وتداخل مع مفاهيم الصدق الحقيقية فمما لا شك فيه أنه إذا بدا الاختبار للفاحصين أو المفحوصين لا معنى له، أو لا علاقة له بالظاهرة موضوع القياس، أوغير مناسب أو ساذجا أو سخيفا. فمن المحتم أن تعاون هؤلاء سيكون ضعيفا حتى ولو كان الصدق الحقيقى للاختبار عاليا، بعبارة أخرى لابد أن يتوافر فى الاختبار هنا الصدق الظاهرى حتى يكون أكثر فاعلية فى مواقف القياس العملية.

ويمكن أن يتحسن الصدق الظاهرى للاختبار بمجرد إعادة صياغة أسئلته بحيث تبدو أكثر ارتباطا بموقف القياس الفعلى، فمثلا إذا كان لدينا اختبار للاستدلال الحسابى البسيط نريد أن نطبقه على مجموعة من المتقدمين للمهن الميكانيكية يحسن أن تستبدل أسئلة الحساب التقليدية التى تتضمنها الكتب المدرسية بأسئلة أخرى أكثر ارتباطا بالعمل الميكانيكى دون أن يحدث أى تغيير يذكر فى الوظيفة السلوكية التى نقيسها، ومع ذلك فإن الصدق الظاهرى ذاته لايمكن أن يكون بديلا لأنواع الصدق الأخرى التى تتحدد بالوسائل الموضوعية، كما أن تحسين الصدق الظاهرى لايؤدى بالضرورة إلى تحسين الصدق الموضوعى.

طرق تحديد صدق المحتوى:

توجد عدة طرق لتحديد صدق المحتوى نذكر منها ما يلى:

١ - **جدول المواصفات:** يساير هذا النوع من الصدق عملية بناء الاختبارات خطوة خطوة، لأن إعداد أسئلة أو مفردات هذه الاختبارات إنما يتم عن طريق الفحص

الدقيق الذى يعتمد على تحليل المحتوى للمثيرات والاستجابات والأهداف (التى فصلناها آنفا)، واستشارة الخبراء واللجوء إلى المحكمين. وعلى أساس المعلومات التى يتم تجميعها يتم تحديد مواصفات الاختبار Test specifications والتى تتضمن الموضوعات التى يجب أن يشملها الاختبار، والاستجابات التى يجب اختبارها، والطريقة الأكثر يسرا فى تحديد مثل هذه المواصفات هى إعداد جدول ثنائى يسمى جدول المواصفات table of specifications أو مخطط blue print، ولأن هذا الجدول أكثر شيوعا فى الاختبارات التحصيلية نعطى مثالا عليه من هذا الميدان إلا أنه لايوجد ما يمنع من استخدامه فى بناء الاختبارات من الأنواع الأخرى.

جدول مواصفات لعينة المحتوى والسلوك (الأهداف): يلعب جدول المواصفات دوراً مهماً فى الاختبارات المرجعة إلى المعيار والاختبارات المرجعة إلى المحك، والفرق بين هذا الجدول فى الحالتين أنه فى النوع الثانى من الاختبارات يشتمل على كل الأصل الإحصائى للعناصر المهمة للمحتوى والسلوك (الأهداف)، أما فى النوع الأول فإن هذا الجدول يمثل عينة من كل منهما.

ويتم إعداد هذا الجدول فى الحالتين عن طريق الفحص الدقيق للمقررات والكتب الدراسية من ناحية، ونواتج التعلم التى يجب اختبارها (فى ضوء الأهداف) من ناحية أخرى، ثم تحديد الأهمية النسبية للموضوعات والأهداف، وعلى هذا الأساس يتم تحديد عدد المفردات التى ترتبط بكل هدف فى كل موضوع، وذلك بإعداد جدول ثنائى (جدول مواصفات) توضع فيه الأهداف التعليمية أفقيا والموضوعات رأسيا، ويعطى الجدول رقم (١١) مثالا على ذلك، ومنه يتضح عدد المفردات التى يشملها الاختبار بالنسبة لكل موضوع من موضوعات المحتوى فى علاقته بكل هدف من الأهداف التعليمية (التى قد تكون معرفية أو وجدانية أومهارية حركية) وقد ملئت خانات الجدول على أساس الوزن النسبى لكل هدف (فى صورة النسب المئوية الموجودة فى السطر الأخير من الجدول)، والوزن النسبى لكل موضوع (فى صورة النسب المئوية الموجودة فى العمود الأخير من الجدول)، وهذه الأوزان النسبية يجب تحديدها مقدما، والطريقة المتاحة حتى الآن لتحديد الوزن النسبى للأهداف هى

استطلاع آراء الخبراء (*)، أما الوزن النسبى للمحتوى فيمكن تحديده موضوعيا فى ضوء الوقت المخصص لتدريسه فعليا، ويحسن أن تكون فى جميع الحالات فى صورة نسبة مئوية.

وقد ملئت خانات الجدول رقم (١١) فى صورة تكرارات متوقعة، ويقترح فؤاد أبو حطب (٩) المعادلة الآتية لتحقيق هذا الغرض.

$$ز = \frac{أ \times ب}{ن}$$

حيث يدل الرمز (ز) على الوزن النسبى للخانة.

(أ) على الوزن النسبى للهدف (العمود) الذى تقع فيه الخانة فى صورة نسبة مئوية.

(ب) على الوزن النسبى لعنصر المحتوى (السطر) الذى تقع فيه الخانة فى صورة نسبة مئوية.

(ن) على المجموع الكلى للأوزان (وهو هنا ١٠٠).

(*) نحب أن ننبه إلى ابتذال طريقة استطلاع آراء الخبراء واللجوء إلى المحكمين فى عدد كبير من الاختبارات التى ظهرت فى السنوات الأخيرة فى مصر وبعض البلاد العربية، إلى الحد الذى يدفعنا إلى الشك فى كثير من بيانات الصدق التى تجمعت منها، والملاحظ أن كثيرا من الباحثين يلجأون إليها للحصول على معلومات شبه امبريقية من أى نوع وبأى سبيل ربما لسهولتها الظاهرة ويسرها النسبى، وفى رأينا أن الباحثين فى حاجة إلى بذل جهد أكبر فى مجال صدق المحتوى لاختباراتهم.

جدول رقم (١١)

جدول الأوزان النسبية لاختبار تحصيل موضوعى لوحدة فى الجغرافيا بمرحلة التعليم الأساسى

أوزان المحتوى فى صورة نسب مئوية ↓	أوزان الأهداف فى صورة نسب مئوية ←				مجموع أوزان المحتوى ↓
	معرفة		فهم تأثير كل عامل من العوامل المختلفة فى تغيير الطقس	مهارة فى تفسير الخرائط الجوية	
	مصطلحات ورموز	حقائق نوعية			
الضغط الجوى	٥	٥	٧	٥	٪ ٢٢
الرياح	٨	٦	١٠	٨	٪ ٣٢
الحرارة	٤	٣	٥	٤	٪ ١٦
الرطوبة والمطر	٥	٤	٦	٥	٪ ٢٠
السحاب	٢	٢	٤	٢	٪ ١٠
مجموع أوزان الأهداف	٪ ٢٤	٪ ٢٠	٪ ٣٢	٪ ٢٤	٪ ١٠٠

وقد حصلنا على القيم الموجودة فى خانات الجدول رقم (١١) من تطبيق هذه المعادلة. فمثلا الخانة الأولى من اليمين (معرفة رموز ومصطلحات الضغط الجوى) حصلنا على الوزن النسبى ٥٪ من ناتج $\frac{٢٤ \times ٢٢}{١٠٠}$ والخانة التى تليها فى السطر نفسه (معرفة الحقائق النوعية للضغط الجوى) من $\frac{٢٠ \times ٢٢}{١٠٠}$ وهى تساوى ٤,٤٠ قربناها إلى ٥٪، ويتطلب الأمر بالطبع بعض التقريب حتى نحصل على المجموع الكلى للأوزان النسبية للأهداف وعناصر المحتوى فى صورة أجزاء من ١٠٠٪.

ويدل الوزن النسبى لكل خانة من خانات الجدول السابق على النسبة المئوية لعدد المفردات التى يجب أن يشملها الاختبار من النوع الذى تدل عليه الخانة. فمثلا إذا كان اختبارنا يتألف من ٥٠ سؤالا فيجب أن يكون ٥٪ من هذه الأسئلة من النوع الذى يقيس معرفة رموز ومصطلحات الضغط الجوى، ١٠٪ من هذه الأسئلة تقيس فهم العوامل المؤثرة فى الرطوبة والمطر، ٨٪ منها يقيس المهارة فى تفسير خرائط الرياح، وهكذا.

ويوضح الجدول رقم (١٢) عدد الأسئلة المتوقع من كل نوعٍ فى كل خانة على أساس أن يكون العدد الكلى لأسئلة الاختبار ٥٠ سؤالا، واعتمادا على الأوزان النسبية التى حصلنا عليها فى جدول (١١)، فمثلا نتوقع أن يتضمن الاختبار ٣ أسئلة من النوع الذى يقيس معرفة رموز مصطلحات الضغط الجوى على أساس أن هذا العدد هو ٥٪ من المجموع الكلى لأسئلة الاختبار أى $\frac{٥٠ \times ٥}{١٠٠}$ = ٢,٥ قربناها إلى ٣، وهكذا.

جدول رقم (١٢)

جدول مواصفات اختبار تحصيلى موضوعى لوحدة فى مادة الجغرافيا بالمرحلة الإعدادية يتألف من ٥٠ سؤالا على أساس الأوزان النسبية فى الجدول رقم (١١)

المحتوى ↓	معرفة		مجموع أوزان المحتوى	مجموع أوزان المحتوى	مجموع أوزان المحتوى
	رموز ومصطلحات	حقائق نوعية			
الضغط الجوى	٣	٢	٤	٣	١٣
الرياح	٤	٣	٥	٤	١٦
الحرارة	٢	١	٢	٢	٧
الرطوبة والمطر	٢	٢	٣	٢	٩
السحاب	١	١	٢	١	٥
المجموع الكلى للأسئلة	١٢	١٠	١٦	١٢	٥٠

ويفيد جدول المواصفات فى بناء الاختبار وتحديد صدق محتواه من خلال ما يلى:

(أ) تحديد عدد المفردات أو الأسئلة التى يتألف منها الاختبار وقد عرضنا طرق تحديد ذلك فيما سبق.

(ب) تحديد نوع المفردات، فالإختيار من متعدد ليس هو النمط الوحيد مثلا الذى يصلح لقياس جميع الأهداف وجميع جوانب موضوعات المحتوى، ويمكن للقارئ الرجوع إلى الباب الرابع من هذا الكتاب، والمخصص للاختبارات التحصيلية ليتعرف على طبيعة كل نوع من أنواع المفردات وما يقيس من عمليات معرفية، مع التنبيه خاصة إلى ما ينشأ عن بعض هذه الأنواع (وخاصة أسئلة البديلين) من أساليب استجابة response set.

(جـ) تحديد قالب المهام task format التى يتألف منها الاختبار. ومن المؤكد أن الاختبارات المتشابهة فى القالب (ألفاظ أو صور أو أعمال عيانية محسوسة) والمختلفة فى المحتوى (ميكانيكا أو الكترونيات مثلا) تكون فيما بينها أكثر ارتباطا من تلك التى تكون متشابهة فى المحتوى (ميكانيكا مثلا) وتعرض فى قوالب مختلفة (ألفاظ أو صور أو أعمال عيانية)، ويقل الارتباط كثيرا بين الاختبارات إذا اختلفت بالطبع فى كل من القالب والمحتوى (٢٤ : ١٤٦)، وتبدو أهمية هذه المسألة إذا علمنا أن قياس مهارة حركية باختبار لفظى يكون أقل صدقا من قياسها باختبار عملى.

٢ - بعض الطرق الأمبريقية لتحديد صدق المحتوى:

وتوجد عدة طرق امبريقية أخرى للتحقق من صدق المحتوى نذكر منها:

(أ) المقارنة بين الدرجات التى يحصل عليها المفحوصون فى الاختبار أو فى صور متكافئة له، قبل التدريب أو العلاج وبعده للتحقق من التحسن الذى يطرأ على الاختبار نتيجة للتدريب، مما يؤكد تعلق relevance ما يقيسه الاختبار ببرنامج التدريب أو التعليم أو العلاج النفسى.

(ب) دراسة أنماط الأخطاء الشائعة فى الاختبار للتأكد من صدق مفتاح الاختبار.

(جـ) تحليل طرق العمل التى يستخدمها المفحوصون وذلك بإعطاء الاختبار فرديا مع توجيههم إلى «التفكير بصوت عال» أثناء حل مشكلات الاختبار.

(د) دراسة أثر بعض العوامل غير المرتبطة بمحتوى الاختبار كالسرعة أو القراءة فى علاقتها بما يقيسه الاختبار، فالسرعة التى لا ترتبط مثلا باختبار يقيس المهارة فى الكتابة على الآلة الكاتبة أو اليسر العددى لاترتبط باختبار آخر يقيس المحصول اللفظى عند الطفل أو قدرته على الاستدلال وحل المشكلات، والقراءة لا ترتبط باختبار يقيس القدرة الكتابية (عند الموظفين الإداريين) وإنما تكفى فى هذه الحالة قراءة جمل من المستوى الذى يستخدم فى الحياة اليومية.

(هـ) تحديد الخصائص اللفظية والكمية لأسئلة الاختبار أو مفرداته، وهذه العملية هى جزء من عملية تحليل المفردات item analysis أثناء بناء الاختبار، وتفيد فى الوقت نفسه - إذا تمت بنجاح - فى تحديد صدق محتوى الاختبار.

ويشمل التحليل الكيفى للأسئلة الحكم عليها فى ضوء فعاليتها من حيث صياغتها وصلاحيتها، وفى هذا الصدد قد يلجأ الباحث إلى مجموعة من المحكمين لإبداء الرأى فى مختلف الجوانب الكيفية للسؤال، أما التحليل الكمى للمفردات فيتضمن على وجه الخصوص تحديد سهولة (أو شيوع) المفردات، وتمييزها وتحديد صدقها وثباتها، وهذه جميعا تفيد فى الحكم على مدى ملاءمة محتوى الاختبار، أو صدق محتواه، ونتناول هذه الجوانب الكمية هنا بشىء من الإيجاز، وعلى القارئ المهتم بالتفاصيل الرجوع إلى فصول (تحليل المفردات) فى كتب القياس النفسى أو الإحصاء النفسى المتخصصة.

يلجأ عدد من معدى الاختبارات إلى التحكيم بواسطة الخبراء للتحقق من صدق كل مفرده وقد قدم كوهن وزملاؤه Cohen etal (٢٢أ: ١٢٨) معادلة للتحقق من صدق المحتوى (Content Validity Ratio) هى :

$$CVR = \frac{Ne-N/2}{N/2}$$

- حيث CVR تشير إلى نسبة صدق المحتوى

- ne تشير إلى العدد الكلى للمحكمين الذين وافقوا على المفردة

– N تشير إلى العدد الكلى للمحكمين

ويعطى كوهن وآخرون الحد الأدنى لمعامل صدق المحتوى بإختلاف عدد المحكمين وهذه المعاملات دالة عند مستوى P=.05

– فعندما يوافق نصف عدد المحكمين على أن المفردة ضرورية ومناسبة للهدف فإن معامل صدق المفردة = صفر.

– وعندما يوافق أكثر من نصف عدد المحكمين على المفردة فإن معامل صدق المفردة يكون موجبا.

– وعندما يكونوا أقل من النصف فإن المعامل يكون سالبا.

مثال ذلك

إذا كان عدد المحكمين (١٠) ووافق على المفردة الأولى (٩) والمفردة الثانية (٥) والمفردة الثالثة (٤) من العدد الكلى للمحكمين فإن محتوى صدق المفردات الثلاثة يكون كالتالى:

المفردة الأولى $CVR = \frac{9-102}{10/2} \quad 0.80$

المفردة الثانية $CVR = \frac{5-102}{10/2} \quad 0.00$

المفردة الثالثة $CVR = \frac{4-102}{10/2} \quad 0.2$

وعلى ذلك فإن الحد الأدنى الذى نقبله من عدد المحكمين عندما يقل عن ٨ لابد أن يوافقوا جميعهم على محتوى المفرده حتى يكون مقابل الصدق ٠,٩٩ .

العدد الكلى للمحكمين	عدد المحكمين الذى نقبله حتى تكون المفردة صادقة
٨	٧
٩	٨
١٠	٨
١١	٩
١٢	٩
١٣	١٠
١٤	١١
١٥	١١
٢٠	١٤
٢٥	١٧
٣٠	٢٠
٣٥	٢٣
٤٠	٢٦

معاملات سهولة (أو شيوع) المفردات:

يحسب معامل سهولة المفردة بصفة عامة بالمعادلة الآتية:

$$\text{س} = \frac{\text{ص}}{\text{ص} + \text{خ}}$$

حيث س : معامل سهولة المفردة

ص : عدد الإجابات الصحيحة

خ : عدد الإجابات الخاطئة.

ويرى فؤاد أبو حطب (١) أن هذه المعادلة يمكن أن تكون أكثر عمومية حين تستخدم لحساب معاملات شيوع المفردات.

$$\text{ش} = \frac{١}{\text{أ} + \text{ب}}$$

حيث ش = معامل شيوع المفردة.

١ = عدد الأفراد الذين أجابوا على المفردة فى الاتجاه (أ) وليكن نعم فى

استخبارات الشخصية أوصواب فى اختبارات القدرات أو ذكروا المفردة فى اختبارات الإنتاج.

ب = عدد الأفراد الذين أجابوا على المفردة فى الاتجاه (ب) وهو عكس الاتجاه (أ).

وقد استخدم فؤاد أبو حطب مفهوم «معامل الشيوع» Commeness Coefficient لأول مرة فى منتصف السبعينيات فى تقنين اختبار رسم الرجل، وفى استفتاء للمشكلات الدراسية عند طلاب المرحلة الثانوية.

دقة مفتاح الاختبار وثبات مفرداته: تصلح لحساب ثبات المفردات أو الأسئلة الطرق المعتادة لحساب الاختبار ككل ومنها إعادة الاختبار إلا أن من أفضل الطرق هنا ما يسمى طريقة الاحتمال المنوالى وهى من اقتراح جتمان ومعادلتها هى:

$$ر = \frac{ن - ١}{ن} \left(ل - \frac{١}{ن} \right)$$

حيث ر = معامل ثبات السؤال.

ن = عدد الاحتمالات الاختبارية للمفردة أو السؤال، ففى أسئلة البديلين ن = ٢ وفى حالة سؤال الاختيار من بين ٤ بدائل تكون ٤ وهكذا.

ل = الاحتمال المنوالى، أى أكبر تكرار نسبى لأى احتمال اختيارى من الاحتمالات التى تحتوى عليها المفردة أو السؤال.

وهذه الطريقة لا ترتبط بالثبات بمعناه المباشر وهى أقرب إلى التحقق من دقة مفتاح الاختبار.

تمييز المفردات: يمكن استخدام الفرق بين النسبتين المئويتين للأفراد الذين أجابوا على المفردة من المستويين الأعلى والأدنى كمؤشر على تمييز المفردة ويرمز له بالرمز (ف) أو (D).

وتوجد عدة طرق أكثر تعقيدا تستخدم مفاهيم المنحنى الاعتدالى لايتسع المقام لتناولها بالتفصيل.

إلا أن ما يجب أن ننبه إليه أن عدداً كبيراً من الاختبارات النفسية لايتطلب شرط

التمييز هذا، وتظل هذه الاختبارات مع ذلك صادقة من حيث محتواها، ولعل أشهر الأمثلة المعاصرة على ذلك الاختبارات المرجعية المحك.

صدق المفردات أو الأسئلة:

لعل أهم بيانات تحليل المفردات التى تتصل مباشرة بصدق محتوى الاختبار ككل ما يسمى المفردات، والتى تشمل أيضًا صدق الاختبارات الفرعية Sub-tests التى يتألف منها الاختبار الكلى.

وتوجد طريقتان لتحديد صدق المفردات أو الأسئلة أو الاختبارات الفرعية هما:

١ – حساب معامل الارتباط بين درجات المفردة أو السؤال أوالاختبار الفرعى ودرجات الاختبار الكلى (والتى تعد فى هذه الحالة المحك الداخلى للصدق) أودرجات أى محك خارجى، ويتحدد نوع معامل الارتباط المحسوب تبعا لطبيعة درجات الجزء (المفردة أو السؤال أو الاختبار الفرعى) من ناحية ودرجات الكل (المحك الداخلى أو المحك الخارجى) من ناحية أخرى.

فإذا كانت درجات الجزء من نوع الكم المنفصل الثنائى أى من نوع الكل أو لا شىء (مثل نعم ـ لا، صفر – ١ ... إلخ) ودرجات المحك الداخلى أو الخارجى من النوع نفسه (مثل نجاح ـ فشل، تفوق ـ تخلف) بحسب معامل الارتباط الرباعى فى هذه الحالة.

أما إذا كانت درجات أحد مكونى العلاقة (الجزء أو المحك) من نوع الكم المنفصل أوالثنائى، ودرجات المكون الآخر من نوع الكم المتصل يحسب فى هذه الحالة معامل الارتباط الثنائى.

أما إذا كانت درجات كل من الجزء (السؤال أو المفردة أو الاختبار الفرعى) والمحك من نوع الكم المتصل يحسب معامل الارتباط التتابعى لكارل بيرسون.

٢ – طريقة المقارنة الطرفية: وذلك بتقسيم درجات المحك الداخلى أو الخارجى إلى مستويين، ويستخدم هنا أسلوب أشبه بطريقة المجموعات المتضادة تتمثل فى انتقاء مجموعتين متطرفتين من المفحوصين فى ضوء درجاتهم الكلية فى الاختبار (المحك الداخلى) أو فى ضوء درجاتهم فى المحك الخارجى ويمكن أن تكون هاتان

المجموعتان المتطرفتان أولئك الذين ينتسبون إلى الربيع الأعلى (المستوى المرتفع) وإلى الربيع الأدنى (المستوى المنخفض) بالنسبة للدرجات الكلية فى المحك، ويرى فلانجان أن نسب مجموعتين متطرفتين يمكن اختيارهما فى ضوء المحك الداخلى خاصة (أى الدرجة فى الاختبارالكلى) هما أولئك الذين يقعون ضمن نسبة ٢٧ ٪ من المفحوصين الحاصلين على الدرجات العليا فيه، والذين يقعون ضمن نسبة ٢٧ ٪ الحاصلين على الدرجات الدنيا.

وبعد تحديد المجموعتين المتضادتين فى ضوء المحك الداخلى أو الخارجى تتم المقارنة بين أداء المفحوصين فيهما على المفردة أو السؤال أو الاختبار الفرعى، ويستخدم فى هذا الغرض إحدى طريقتين...

أ- فى حالة استخدام نسبة ٢٧ ٪ العليا، ٢٧ ٪ الدنيا يمكن الاعتماد على معاملات الارتباط التى حسبها فلانا جان فى جداوله الشهيرة والتى تعبر عن العلاقة بين معاملات سهولة السؤال أوالمفردة للمجموعتين العليا والدنيا من ناحية ودرجات المفحوصين فى كل من المجموعتين فى السؤال (وهى من نوع الكم المتصل أو الثنائى، وبعبارة أخرى تعبر عن العلاقة بين الدرجة الكلية فى الاختبار ودرجة كل سؤال على حدة من أسئلة الاختبار فى المحك الداخلى أو الخارجى.

ب - فى حالة اختيار المجموعتين المتضادتين فى المحك الداخلى أو الخارجى فأى طريقة أخرى غير الطريقة السابقة (كطريقة الربيع الأعلى والأدنى، أو نسبة ١٠ ٪ الدنيا مثلا) يمكن حساب دلالة الفروق بين أداء المجموعتين فى المفردة أو السؤال أوالاختبار الفرعى باستخدام أحد مقاييس الدلالة الإحصائية المناسب كا٢ أواختبار «ت» أوتحليل التباين أو غير ذلك من الطرق البارامترية واللابارامترية).

ونحب أن ننبه هنا إلى خطأ شاع فى السنوات الأخيرة فى بعض رسائل الماجستير والدكتوراه، بل وفى بعض «البحوث» التى يتقدم بها أصحابها إلى اللجان العلمية الدائمة للترقية إلى وظائف أساتذة وأساتذة مساعدين، ويتلخص هذا الخطأ فى أن الباحث يختار مجموعتيه المتضادتين فى ضوء المحك ثم يقارن بينهما مستخدما درجات المحك ذاتها، وهذا التحليل قد يفيد فى تدعيم فكرة «تضاد» المجموعتين فحسب، ولايتجاوز الأمر إلى أبعد من ذلك، ولايفيد على أى نحو - بالطبع- فى تقديم

أى معلومات عن صدق الاختبار، فالصدق يتطلب - ونكرر - أن يتم تحليل أداء المجموعتين المتضادتين فى المحك على المفردة أو السؤال أو الاختبار الفرعى.

ويمتد أصحاب الخطأ السابق أحيانا إلى ما هو أخطر، فنجدهم يكونون المجموعتين المتضادتين فى ضوء أداء الأفراد فى الاختبار الفرعى، وليس فى المحك، ثم يقارنون بين المجموعتين فى درجاتهما فى الاختبار الفرعى نفسه، باستخدام مقاييس الدلالة الإحصائية، وبالطبع يحصلون على فروق دالة يستنتجون منها «صدق» الاختبار الفرعى، أليس هو العبث العلمى بعينه، هو نوع من تحصيل الحاصل، الذى لا يضيف جديدا إلى صدق الاختبار، ولايتجاوز حدود «اللعب الإحصائى» السخيف.

(٣) الصدق المرتبط بالمحكات

يدل الصدق المرتبط بالمحكات criterion-related validity على قدرة الاختبار على التنبؤ بسلوك المفحوص فى مواقف محددة أو تشخيص هذا السلوك، ولهذا لابد من الحكم على الأداء فى الاختبارات فى ضوء أحد المحكات، ويقصد بالمحك مقياس مباشر ومستقل لما يهدف الاختبار إلى قياسه والتنبؤ به أو تشخيصه، أو هو ميزان لتحديد مدى صلاحية الاختبار، أو هو اختبار للاختبار، وعلى ذلك فبالنسبة لاختبار يقيس الاستعداد الميكانيكى يمكن أن يكون المحك أداء المفحوصين فى عمل ميكانيكى فعلى فى مدرسة ثانوية صناعية أو مصنع، وهنا يجب أن نميز بين غرضين يستخدم فيهما هذا النوع من الصدق: أولهما، التنبؤ الطويل المدى، وثانيهما التشخيص، وذلك فى ضوء العلاقات الزمنية بين المحك والاختبار، فإذا تصاحبت أو تلازمت زمنيا بيانات المحك ودرجات الاختبار يصبح الصدق فى هذه الحالة من النوع «التلازمى»، أما إذا وجد فاصل زمنى طويل (قد يصل لعدة سنوات) بين درجات الاختبار ومعلومات المحك فإن الصدق يصبح «تنبؤيا». وهكذا يمكن أن نميز بين نوعى الصدق هذين فى ضوء أهداف القياس، فالمعلومات التى يوفرها الصدق التنبؤى ترتبط بالاختبارات التى تستخدم فى انتقاء الأفراد وتصنيفهم وتوجيههم تربويا أو مهنيا أو عسكريا وكذلك فى أغراض التنبؤ الكلينيكى، وفى جميع هذه الأغراض تتشابه إجراءات الصدق، حيث نعطى الاختبار لعينة من الأفراد تمثل المجتمع الأصلى

(موضوع الاهتمام)، ولكن الدرجات التى يحصل عليها أفراد هذه العينة لا تستخدم فى اتخاذ أية قرارات خاصة بهم، كما يجب ألا تعطى هذه الدرجات إلى أى شخص أو سلطة يمكن له أو لها أن تتخذ قرارات تؤثر فيهم، حتى لا يطرأ على المحك أى تشويه أو خلط أو تلوث criterion contamination لأن الدرجة الاختبارية فى هذه الحالة قد تؤثر على الوضع النسبى للمفحوص فى المحك، وخاصة إذا كان من نوع التقديرات، وبعد ذلك تتم متابعة العينة الأصلية حتى تتوافر لدينا عنها المعلومات التى نستقيها من المحك والتى تؤدى بنا إلى تحديد المدى الذى تتفق فيه درجات الاختبار مع النتائج التى نلاحظها كما تتمثل فى درجات المحك، وهذا الاتفاق يحدد لنا مدى تنبؤية الاختبار، وتستخدم لهذا الغرض أساليب إحصائية معقدة تعتمد فى جوهرها على أسلوب تحليل الانحدار regression analysis (راجع ١).

وعادة ما يتضمن دليل الاختبار قيما خاصة بالصدق التنبؤى مشتقة من استخدام محكات مختلفة بهدف مساعدة من يستخدم الاختبار فى فهم ما يقيسه وذلك على الرغم من أن مستخدمى الاختبارات قد لا يهتمون اهتماما مباشرا بالتنبؤ بأى من المحكات المستخدمة.

أما الصدق التلازمى فهو أكثر ملاءمة للاختبارات التى تستخدم لأغراض التشخيص(*) لا التنبؤ بنتائج المستقبل، وفى رأينا أن الغرض الأخير هو أكثر الأغراض مشروعية فى إطار المرحلة الراهنة من تطور علم القياس النفسى.

ويلجأ الباحثون إلى الصدق التلازمى فى بعض الأحيان كنوع من حلول الوسط لمشكلة الصدق التنبؤى وما تتطلبه من وقت طويل، وفيها تعطى الاختبارات لمجموعة من الأفراد تتوافر لدى الباحث عنهم معلومات خاصة بمحك معين، فتقارن مثلا درجات طلاب الجامعة فى اختبار معين بمتوسط تراكمى لدرجاتهم الدراسية حتى وقت الاختبار كما نقارن درجات العمال فى الاختبار بمستوى نجاحهم الراهن فى العمل.

وتتمثل أهم مشكلات الصدق التلازمى فى صعوبة تحديد علاقات السبب والأثر،

(*) يستخدم فؤاد أبو حطب (١) مصطلح بمعنى أكثر شمولا من معناه الشائع فى إطار علم النفس الكلينيكى. ويقصد به هنا وصف الوضع الراهن للسلوك وقد يمتد إلى الأحداث الماضية.

التقويم النفسى ـــــــ

ميزان للاختبار أو اختبار للاختبار، ولهذا لابد من التحوط والحذر فى انتقاء المحكات حتى لا توزن الاختبارات بموازين فاسدة أو تختبر باختبارات مضللة، ونعرض فيما يلى لأنواع المحكات الشائعة.

١ - **التحصيل المدرسى الأكاديمى العام**: هو أشهر المحكات وخاصة بالنسبة لاختبارات الذكاء، ولذلك توصف هذه الاختبارات عادة بأنها مقاييس الاستعداد المدرسى، ومن مؤشرات هذا التحصيل المدرسى مجموع درجات الاختبارات التحصيلية، ونتائج النقل وتقديرات التخرج، ومراتب الشرف، والجوائز أو المنح، وتقديرات المعلمين لتحصيل تلاميذهم، ويستخدم هذا المحك - إلى جانب اختبارات الذكاء كما أشرنا - فى حساب صدق مقاييس الشخصية، سواء كان غرضنا من استخدام هذه الإختبارات هو التنبؤ أو التشخيص، وهو محك شائع فى جميع المراحل التعليمية وفى جميع أنواع التعليم..

٢ - **الأداء فى برنامج معين**: هذا المحك أكثر ارتباطا بحساب صدق بطاريات الاستعدادات الفارقة واختبارات القدرات المتخصصة، ومقاييس بعض سمات الشخصية، وهو يعتمد على مؤشرات أداء المفحوصين فى برنامج معين من التدريب المتخصص، ففى حساب صدق اختبار معين للقدرة الميكانيكية يمكن الاعتماد على التحصيل النهائى فى مقرر معين يتعلق بإحدى المهن الصناعية، وكما يمكن أن نعتبر التحصيل فى مقررات التعليم بإحدى المهن الصناعية. وكما يمكن أن نعتبر التحصيل فى مقررات التعليم التجارى، مثل الاختزال والآلة الكاتبة ومسك الدفاتر، محكات لاختبارات الاستعدادات فى هذه الميادين، وكذلك فإن أداء طلاب المعاهد الموسيقية والفنية قد يعتبر محكا لحساب صدق اختبارات الإستعدادات الموسيقية والفنية وقد حسب صدق معظم اختبارات الاستعدادات المهنية باستخدام محكات التحصيل فى كليات الحقوق والطب والصيدلة والهندسة وغيرها، أما إذا كانت الاختبارات من النوع الذى نستخدمه فى أغراض برنامج معين للانتقاء، فقد نستخدم سجلات التدريب كمصدر للمعلومات التى نحتاج إليها، كما هو الحال فى حساب صدق اختبارات انتقاء الطيارين، فالمحك هنا هو أداء الطيارين فى مواقف التدريب الفعلية.

ومن المؤشرات الخاصة التى يتضمنها محك الأداء فى برامج التدريب يمكن أن

نذكر الاختبارات التحصيلية المتخصصة التى نطبقها عقب الانتهاء من برنامج التدريب أوتقديرات المدربين أو النجاح والفشل فى التدريب (الاستمرار والإكمال فى مقابل الانقطاع والاستبعاد)، وقد نستخدم درجات الامتحانات فى مقررات بذاتها فى المدرسة الثانوية أوالكليات الجامعية والمعاهد العليا وذلك لتحديد صدق بطاريات الاستعدادات، فمثلا يمكن مقارنة درجات الطلاب فى اختبارات الفهم اللغوى بدرجاتهم فى مقررات اللغة العربية، ودرجات التصور البصرى والمكانى بدرجات مقرر الهندسة... إلخ.

ولابد أن نميز فى هذا الصدد بين ما يسمى المحكات المتوسطة والمحكات القصوى أو النهائية (١٣) بالنسبة لمؤشرات سجلات التدريب، فعند إعداد اختبار لقياس الاستعداد للطيران أو لقياس الاستعداد الطبى أو الاستعداد للتدريس مثلا نجد أن المحكات النهائية هى الأداء الفعلى فى ميدان القتال، أو النجاح فى ممارسة الطب، أوالنجاح فى الممارسة الفعلية للتدريس داخل حجرة الدراسة، ويتطلب ذلك دون شك وقتا طويلا حتى يمكننا الحصول على معلومات من هذه المحكات، بل إنه من المشكوك فيه إمكان وصولنا إلى محك نهائى فى الواقع، بالإضافة إلى أنه، حتى إذا توفر لنا مثل هذا المحك النهائى، فإنه قد يتعرض لعوامل كثيرة يصعب التحكم فيها مما يجعل المحك فى النهاية لا فائدة منه. فمن الصعب مثلا أن تقوّم درجة النجاح النسبى للأطباء الذين يمارسون مختلف التخصصات الطبية فى أجزاء مختلفة من الوطن سواء داخل المستشفيات أو عيادتهم الخاصة، ولهذا السبب نلجأ فى العادة إلى المحكات المتوسطة التى ذكرنا أمثلة لها فيما سبق مثل سجلات الأداء فى مرحلة معينة من مراحل التدريب.

٣ – الأداء على العمل نفسه: هو أكثر المحكات صلاحية، لأنه يعتمد على سجلات تتبعية للأداء الواقعى والفعلى للأفراد فى العمل، وقد استخدم العلماء هذا المحك إلى حد ما فى حساب صدق اختبارات الذكاء العام والشخصية، وإلى حد كبير فى حساب صدق اختبارات الاستعدادات الفارقة والقدرات المتخصصة، ومن المؤكد أن «الأعمال» تختلف فى مستوياتها وفى أنواعها، ويصدق هذا على الأعمال الصناعية والتجارية والمهنية والعسكرية وغيرها. ورغم أن معظم مقاييس أداء «العمل» قد لا تعتبر محكات نهائية إلا أنها أفضل بكثير من محك سجلات التدريب الذى تحثنا عنه

آنفا، ومع ذلك فقد لا يسمح قياس أداءالعمل بتوحيد وتقنين الشروط التى تتوافر أثناء التدريب ، كما أنه يتطلب تتبعا طويل المدى مما يؤدى إلى قلة عدد المفحوصين، وبالتالى يؤثر فى النتائج على نحو لا حيلة للباحث فيه.

والمؤشرات التى نستخدمها فى قياس درجة النجاح فى «العمل» عديدة منها مقدار الإنتاج، ونوعه، والاستهداف للحوادث، والمرتبات والعلاوات والترقيات والاستمرار فى العمل وعدم الانتقال إلى غيره، وطول مدة الخدمة ومعدل النمو فى المهنة، وتقديرات المشرفين والرؤساء للكفاءة المهنية للمفحوص وقد يعتمد مقياس المحك على ملاحظة عينة صغيرة محددة من أداء الشخص فى العمل، أو قد يشتق من سجل تتبعى للإنتاج، أو تاريخ الشخص فى العمل، بحيث يصبح محكا شاملا للفترة التى يقضيها المفحوص فى عمله.

٤ – **المجموعات المتضادة:** محك «المجموعات المتضادة» Contrasted groups، هو من نوع المحكات المركبة التى تدل على الآثار المتراكمة وغير المضبوطة للحياة اليومية، أى أنه أقل وضوحا وأكثر تعقيدا من المحكات الأخرى التى سبق مناقشتها، ويعتمد فى أساسه على بقاء الفرد واستمراره داخل مجموعة معينة لفترة طويلة نسبيا من الزمن أوعدم بقائه فيها وخروجه منها. ففى حسابنا لصدق اختبار الذكاء يمكن أن نقارن درجات ضعاف العقول المقيمين بإحدى المؤسسات بالدرجات التى يحصل عليها تلاميذ مدرسة عادية فى نفس السن تقريبا، وفى مثل هذه الحالة يتأثر المحك بالعوامل العديدة التى تتحكم فى هذه المؤسسات بوجه عام وتجعلها بيئة مختلفة عن المدرسة العادية، وكذلك إذا أردنا حساب صدق اختبار للاستعداد الموسيقى أو الاستعداد الميكانيكى فيمكن أن نقارن بين الدرجات التى يحصل عليها الطلاب الذين يدرسون الموسيقى أو الهندسة فى كليتين أو معهدين متخصصين، وبين درجات طلاب آخرين يدرسون فى كلية جامعية لا تعد طلابها إعدادا منظما فى هذين المجالين.

ويمكن للباحث الذى يلجأ إلى هذه الطريقة فى حساب الصدق أن يكون المجموعات المتضادة مستخدما ما يشاء من المحكات، مثل الدرجات المدرسية أو تقديرات النجاح والفشل أو الأداء فى العمل أو غيرها، وذلك باختيار الحالات المتطرفة

فى توزيع درجات المحك، وتصبح المجموعات المتضادة فى هذه الحالة هى جماعات تمايزت تدريجيا بفعل مطالب الحياة اليومية وخبراتها العديدة.

وتستخدم طريقة المجموعات المتضادة فى حساب صدق اختبارات الشخصية على وجه الخصوص، فمثلا يمكن أن نقارن أداء من يمارسون أعمالا قيادية أو اقناعية بأداء من يمارسون أعمالا كتابية فى اختبار معين للسمات الاجتماعية ليمكن حساب صدق هذا الإختبار مفترضين أن الأفراد الذين يعملون بمهن معينة تتطلب التعامل الاجتماعى (البيع - التدريس ... إلخ) . وظلوا يقومون بها لفترة من الزمن سوف يتقدمون على غيرهم من الأفراد الذين يعملون فى ميادين أخرى لا تتطلب مثل هذا النوع من التعامل كالأعمال الكتابية والهندسية، بل يمكن أن نقارن بين أداء الطلاب الذين يقومون بنشاط كبير خلال فترة دراسية معينة، كما أننا كثيرا ما نلجأ إلى الجماعات المهنية لحساب صدق اختبارات الميول (ومنها الميول المهنية) ومقاييس الاتجاهات، ومن الجماعات التى تستخدم فى حساب صدق مقاييس الجماعات السياسية أو الدينية أو غيرها من الجماعات الخاصة التى تمثل وجهات نظر مختلفة اختلافا واضحا حول موضوعات اجتماعية معينة، وقد تستخدم جماعات معينة مثل الأحداث المنحرفين المودعين فى المؤسسات فى مقابل الأحداث غير المنحرفين، أو العصابيين فى مقابل الأسوياء، وعند حساب صدق اختبارات الشخصية التى تتعلق بالتكيف الانفعالى أو الاجتماعى، وفى بعض هذه الاختبارات يلجأ الباحثون إلى التشخيص الطبى النفسى باعتباره أساسا لاختيار الأسئلة ودليلا على صدق الاختيار، إلا أن التشخيص الطبى النفسى لابد أن تتوفر عدة شروط مهمة ليصبح محكا مفيداً، من أهمها أن يستند إلى ملاحظة طويلة المدى ودراسة مفصلة لتاريخ الحالة، أما إذا اعتمد على مجرد مقابلة شخصية قصيرة أو على فحص طبى نفسى غير عميق فلن يكون هناك سبب يدعونا إلى توقع أن يكون مثل هذا التشخيص أفضل من الدرجة التى يحصل عليها الفرد فى الاختبار موضع البحث، بعبارة أخرى يصبح المحك المفترض (التشخيص الطبى النفسى) كالاختبار الجديد نفسه، مجرد مؤشر يحتاج هو فى ذاته إلى مزيد من البحث حول صدقه ولايكون محكا كافيا نعتمد عليه فى تقدير صدق الاختبار.

٥ - **التقديرات**: سبق أن تحدثنا عن التقديرات ratings التى يعطيها المعلمون

لتلاميذهم أو المشرفون على العمال لعمالهم وغيرهم عند حديثنا عن المحكات الأخرى والتى نطلبها بهدف الحصول على معلومات تتصل بهذ المحكات مثل التحصيل المدرسى الأكاديمى أو الأداء فى نوع متخصص من الأعمال أو النجاح المهنى، إلا أن التقديرات قد تصبح محكا نعتمد عليه فى ذاته عند حساب الصدق حين تتطلب حكما شخصيا من جانب الملاحظ (المعلم أو المشرف) على الخاصية النفسية التى يقيسها الاختبار، ومن ذلك أن نطلب من الملاحظين (المعلمين أو المشرفين) أن يعطوا تقديرات للمفحوصين فى خصائص معينة تقيسها الاختبارات مثل السيطرة أو القيادة أو الأصالة أو الابتكار أو الأمانة... إلخ. هذه التقديرات تعد ذاتها محكات واستخدمها الباحثون فى حسابهم لصدق كثير من الاختبارات، وخاصة اختبارات الشخصية وذلك لصعوبة الحصول على محكات موضوعية فى هذا الميدان وخاصة بالنسبة للسمات الاجتماعية التى لايمكن الحكم عليها إلا من خلال مثل هذه التقديرات المعتمدة على العلاقة الشخصية.

والواقع أن التقديرات التى نحصل عليها من ملاحظين لديهم التدريب والخبرة وفى ظروف مضبوطة ضبطا تجريبيا دقيقا توفر لنا معلومات مهمة عن المحك. ومما يزيد ثقتنا فى هذا المحك أن نحصل على تقديرات مستقلة من أكثر من ملاحظ وحساب معاملات الموضوعية (راجع الفصل الرابع) حتى يمكن التغلب على العوامل الذاتية التى تؤثر فى ثبات هذه التقديرات. كما يمكن أن تزداد دقة التقديرات إذا استخدمنا مقاييس التقدير rating scales بحيث تتضمن وحدات محددة تحديدا واضحا مع وجود ضمانات خاصة للتخلص من أخطاء التقدير الشائعة مثل أثر وقع الهالة halo effect (وهى عبارة عن ميل الملاحظين أوالفاحصين إلى التأثر بسمة أو سمات غير مطلوبة، وتكون فى صالح المفحوصين أو فى غير صالحة وتؤدى إلى التأثير فى أحكامام وتقديراتهم للسمة موضع القياس) trait acquaintance وفهم لمعناها بحيث يمكنهم الحكم على وجودها بدرجة معينة لدى المفحوصين، بعبارة أخرى لايكفى أن يعرف الفاحص المفحوص فترة طويلة ليصبح قادرا على إصدار حكم عليه وإنما يجب أن تكون قد تهيأت له فرصة ملاحظته فى مواقف مختلفة تظهر فيها السمة السلوكية موضوع الاهتمام، وفى جميع الحالات لايجب أن يعطى الفاحصون تقديرات للمفحوص فى سمة ليس لديهم عنها معرفة كافية.

وتلعب التقديرات دوراً مهماً فى مقاييس المحك الذى يعتمد على التقويم الكلينيكى. وفى مثل هذه الأحوال لابد من تدعيم الحكم الكلينيكى ببعض المعلومات الإضافية مثل التاريخ المفصل للحالة ودرجات الاختبارات الأخرى وغيرها من الوسائل الموضوعية، أما التقويم النهائى فيعتمد على عملية الحكم كما هو الحال فى جميع التقديرات.

٦ - معاملات الارتباط بين الاختبار وغيره من الاختبارات: من وسائل تحديد صدق الاختبار حساب معاملات الارتباط بين درجات الاختبار التجريبى (الجديد) وبين درجات اختبارات أخرى قائمة بالفعل تقيس ما يفترض الاختبار الجديد قياسه. وقد نلجأ إلى هذه الطريقة إذا كان الاختبار الجديد صورة مختصرة أو مبسطة لاختبار موجود فعلا، وفى هذه الحالة يكون الاختبار الأصلى محكا، كما يمكننا أن نقدر صدق اختبار لفظى (ورقة وقلم) باستخدام اختبار يدوى عملى مطول يستغرق وقتا طويلا وله معامل صدق مرتفع تم تحديده من قبل، وكذلك يمكن تقدير صدق اختبار جماعى باستخدام اختبار فردى (مثل اختبار ستانفورد - بينيه الذى استخدم كثيرا كمحك لحساب صدق الاختبارات الجماعية للذكاء). وفى مثل هذه الأحوال يمكن اعتبار الاختبار أنه ما لم يكن الاختبار الجديد بديلا مبسطا أو مختصرا للاختبار الأصلى يصبح لا قيمة له، لأنه لن يكون فى أحسن حالاته إلا تكرارا لاختبار موجود بالفعل.

بعض مشكلات الصدق المرتبط بالمحكات: لعل من أهم المشكلات التى تعترض نظرية القياس النفسى بالنسبة للصدق المرتبط بالمحكات ما يتصل بمدى عمومية هذا النوع من الصدق، وخاصة أن أهم ما يميزه أنه نوع من الصدق العاملى للاختبار فى موقف بذاته، كأن يحدد فعالية الاختبار لبرنامج تدريبى معين.

وقد بدأت المشكلات فى الظهور حين لوحظ أن معاملات الارتباط بين اختبارات الاستعداد المقننة والأداء فى أعمال يفترض فيها أن تتشابه مع ما تقيسه هذه الاختبارات، وخاصة فى بحوث الصدق فى الأعمال الصناعية لوحظ أن معاملات الصدق اختلفت اختلافات بينه، كما حصل الباحثون على ما يشبه ذلك فى بحوث الصدق التى استخدمت التحصيل فى مختلف المقررات الدراسية محكا، وأدى ذلك إلى

سيطرة اتجاه لدى مستخدمى الاختبارات حتى منتصف السبعينيات خلاصته أن صدق الاختبار لا يقبل التعميم إلى مواقف مختلفة، وسادت فكرة «الخصوصية الموقفية»، واعتبرت فى الوقت نفسه عائقا لاستخدام الاختبارات المقننة فى انتقاء الأفراد. إلا أن الدراسات السيكومترية التى أجريت طوال العقدين الماضيين أكدت أن هذه الاختلافات فى معاملات الصدق المرتبط بالمحكات هى «محض اصطناع إحصائى» ناتج عن صغر حجم العينات المستخدمة فى حساب الصدق، وعدم ثبات المحك، وضيق مدى الفروق الفردية فى عينات التقنين، وهذه جميعا تؤدى إلى اختلاف معاملات الصدق المستخدمة فى حساب الصدق، وعدم ثبات المحك، وضيق مدى الفروق الفردية فى عينات التقنين، وهذه جميعا تؤدى إلى اختلاف معاملات الارتباط من ناحية، وإلى خفضها من ناحية أخرى.

وقد أمكن التغلب على هذه المشكلة فى السنوات الأخيرة من خلال نموذج جديد اقترحه شمدت Schmidt وهانتر Hunter وأطلقا عليه اسم «النموذج البايزى لتعميم الصدق» Bayesian Validity Generalization Model وطبقاه على عينات كثيرة اختيرت من تخصصات مهنية كثيرة وأكد أن صدق اختبارات الاستعدادات اللفظية والعددية والاستدلالية يمكن أن يعمم إلى عدد كبير من المهن، وأن الاختلافات السابقة لم تكن إلا محض «مصادفة».

(جـ) صدق التكوين الفرضى:

يقصد بصدق التكوين الفرضى Construct validity مدى قياس الاختبار لتكوين فرضى أو مفهوم نفسى أن سمة، ومن أمثلة هذه التكوينات الفرضية الذكاء والفهم الميكانيكى والطلاقة اللغوية والعصابية والقلق وغيرها. وكل من هذه المفاهيم أوالتكوينات يستنتج من أساليب الأداء كما يفسر حدوث الاتساق الملحوظ فى الاستجابة (٢٤). ويعتمد هذا النوع من الصدق على وصف أوسع ويتطلب معلومات أكثر حول الخاصية السلوكية موضوع القياس نحصل عليها من مصادر مختلفة، وأي معلومات تفيد فى إلقاء الضوء على المفهوم الذى نحاول قياسه وعلى العوامل التى تؤثر تضاف إلى رصيد صدق التكوين الفرضى.

وتوجد أساليب كثيرة يمكن أن تسهم فى تحديد هذا النوع من الصدق وهى ما يأتى:

١ – تمايز العمر والتغيرات الارتقائية: يمكن أن نلجأ إلى هذه الطريقة فى حسابنا لصدق اختبارات ذكاء الأطفال (كاختبار ستانفورد ـ بينيه، أو معظم اختبارات الذكاء لأطفال ما قبل المدرسة)، وذلك لتحديد ما إذا كانت الدرجات تتزايد بتقدم العمر مفترضين أن القدرات تنمو مع العمر خلال الطفولة وبالتالى نتوقع زيادة فى درجات الاختبار إذا كان الاختبار صادقا، ويعود الفضل إلى بينيه فى صياغة هذا المفهوم الخاص بتحديد المقياس العمرى للذكاء.

ومن الواضح أن محك تمايز العمر لايمكن استخدامه فى الوظائف السلوكية التى لا تظهر تغيرات مطردة مع تقدم العمر ومن ذلك مقاييس الشخصية بوجه عام، كما أننا يجب أن نشير إلى أن هذا المحك محك سلبى وليس محكا إيجابيا، بعبارة أخرى إذا لم تزد درجات الاختبار مع تزايد العمر فإن ذلك قد يدل على أن الاختبار غير صادق فى قياسه للقدرات التى يدعى أنه عينة لها، إلا أن تحسين الاختبار بحيث يصبح مقياسا متزايداً مع تزايد العمر لا يحدد لنا فى ذاته الميدان الذى يقيسه الاختبار، فمثلا يظهر كل من الطول والوزن زيادة منتظمة مع تزايد العمر ومع ذلك لايمكن أن نعتبر أحدهما أو كليهما من مقاييس الذكاء، وكذلك يجب أن نؤكد أنه فى الاختبار النفسى الذى يقيس خصائص سلوكية تتزايد مع العمر إنما تتأثر هذه الزيادة أوالنمو بظروف البيئة التى يقدن فيها الاختبار، فالظروف البيئية والحضارية قد تستثير بعض الخصائص السلوكية وقد تعطل البعض الآخر، ومعنى ذلك أن محك تمايز العمر لايمكن أن نعتبره محكا شموليا بحال من الأحوال، وإنما هو كغيره من المحكات محكوم بالموقف الحضارى الخاص الذى نشتقه منه.

٢ – معاملات الارتباط والاختبارات الأخرى: وقد نلجأ فى تقدير صدق التكوين الفرضى إلى حساب معاملات الارتباط بين الاختبار التجريبى وبين الاختبارات التى تتشابه معه نظريا فى العمليات التى يقيسها، ويسمى كامبل Campbell وفيسك Fiske الصدق التقاربى converant validity ومن أمثلة ذلك معامل ارتباط اختبار الاستدلال الكمى بدرجات لمفحوصين فى اختبار التحصيل الرياضى، إلا أن هذا النوع

من معاملات الارتباط لايكفى ولذلك يقترح كامبل وفسك ضرورة أن يضاف إلى ذلك حساب معاملات الارتباط بين الاختبار واختبارات أخرى يفترض نظريا أنها تختلف عنه، ويسميان هذا النوع من الصدق باسم الصدق التمييزى discriminant validity.

ويستخدم الصدق التمييزى لمعرفة مدى تحرر الاختبار التجريبى من أثر العوامل التى لا علاقة لها بالعملية السلوكية موضع القياس، فمثلا إذا حصلنا على معامل ارتباط صفرى بين اختبار معين للاستعدادات واختبارات الفهم اللغوى فإن ذلك قد يعنى أن هذا الفهم اللغوى لايرتبط بالاستعداد موضوع القياس فى الاختبار الجديد وكثيرا ما تتضمن كتيبات الاختبارات معاملات الارتباط المنخفضة بين الاختبار التجريبى واختبارات أخرى تقيس عمليات مختلفة وهى مؤشرات على الرغم من أنها غير مباشرة وسلبية إلا أنها تدل فى الوقت نفسه على صدق الاختبار إذ تعنى أن الاختبار يقيس عمليات أخرى لم توضع فى الحسبان. إلا أن معاملات الارتباط المنخفضة هذه قد لا تعنى فى ذاتها أن الاختبار التجريبى صادقا، ولكى تستخدم طريقة الصدق التقاربى والصدق التمييزى بكفاءة يقترح كامبل وفسك تصميما تجريبيا خاصًا يسمياه «مصفوفة السمات المتعددة والطرق المتعددة» Multitrait-Multimethod Matrix.

ويتطلب هذا التصميم تقويم سمتين نفسيتين أو أكثر بطريقتين مختلفتين أو أكثر، ثم تحسب معاملات الارتباط الممكنة بين درجات الاختبارات التى تقيس السمات موضع البحث، ولكى نوضح هذه الطريقة نعطى المثال الافتراضى التالى الذى يقترحه الباحثان.

نفرض أننا أردنا قياس ثلاث سمات هى (أ) السيطرة، (ب) الاجتماعية، (جـ) دافعية الإنجاز، بثلاث طرق مختلفة هى (١) قائمة من نوع التقرير الذاتى، (٢) أسلوب إسقاطى، (٣) تقديرات الأقران والزملاء، فإننا نشير فى هذه الحالة إلى سمة السيطرة كما تقاس بالتقرير الذاتى بالرمز (أ١) وكما تقاس بالاختبار الإسقاطى (أ٢)، وإلى سمة دافعية الإنجاز كما تقاس بتقديرات الأقران بالرمز (جـ٣)، وهكذا بالنسبة لجميع البدائل الممكنة.

ويوضح الجدول رقم (١٣) مصفوفة ارتباط جميع الدرجات التى يمكن الحصول عليها بهذه الطريقة، وقد وضعت معاملات الثبات الخاصة بطريقة قياس كل سمة بين قوسين.

وفى هذا الجدول رقم (١٣) نجد (١) معاملات الصدق فى صورة علاقات بين الدرجات التى حصل عليها المفحوصون لنفس السمة بطرق مختلفة مستقلة (كما هو مفترض فى إجراء الصدق المعتادة)، (٢) معاملات الارتباط بين مختلف السمات التى قيست بنفس الطريقة، و(٣) معاملات الارتباط بين السمات المختلفة التى قيست بطرق مختلفة، ولكى يحصل الباحث على صدق التكوين الفرضى لابد بالطبع أن تكون معاملات الصدق (النوع الأول) أعلى من النوعين الآخرين من معاملات الارتباط. فنحن نتوقع مثلا أن معامل الارتباط بين دافعية الإنجاز كما تقاس بوسائل التقرير الذاتى وكما تقاس بالاختبار الإسقاطى يكون أعلى من معامل الارتباط بين دافعية الإنجاز والسيطرة كما تقاس بوسائل التقرير الذاتى، فإذا حدث أن المعامل الأخير كان مرتفعا فإننا قد نستنتج من ذلك أن أداء الفرد فى الاستخبار أوالقائمة المستخدمة يتأثر بعامل مشترك لايرتبط بما يقيسه (أى نوع من تباين الخطأ الذى تناولناه بالتفصيل فى الفصل الرابع المخصص لموضوع الثبات)، وقد يكون هذا العامل متصلا بالفهم اللغوى لأسئلة أو العبارات، أو رغبة المفحوص فى أن يظهر بصورة موجبة فى جميع السمات (المرغوبية الاجتماعية).

وتذكر أنستازى (١٣ : ١٤٩) أن معامل الثبات فى مصفوفة السمات المتعددة والطرق المتعددة يمثل درجة الاتفاق بن مقياسين لنفس السمة بينهما أقصى حد من التشابه (كما هو الحال فى الصور المتكافئة لنفس الاختبار). أما الصدق فيمثل درجة الاتفاق بين مقياسين لنفس السمة بينهما أقصى حد من الاختلاف (كما هو الحال ففى درجات الاختبار وتقديرات المعلمين أوالرؤساء). وطالما أن التشابه والاختلاف بين الطرق هو مسألة نسبية فإننا - من الوجهة النظرية البحتة - يمكن أن نعتبر الثبات والصدق إذا ذكرنا أن الثبات فى موجز يهتم بتباين الخطأ، بينما الصدق يكاد يتطابق مع مفهوم التباين الحقيقى، وكلاهما يؤلف التباين الكلى أو الفروق الفردية فى الدرجات التى يحصل عليها المفحوصون فى الاختبار.

جدول رقم (١٣)
مصفوفة ارتباط متعددة وطرق متعددة
(عن ١٣ : ١٤٨)

السمات		الطريقة ١			الطريقة ٢			الطريقة ٣		
		أ١	ب١	جـ١	أ٢	ب٢	جـ٢	أ٣	ب٣	جـ٣
الطريقة ١	أ١	(٠,٨٩)								
	ب١	٠,٥١	(٠,٨٩)							
	جـ١	٠,٣٨	٠,٣٧	(٠,٧٦)						
الطريقة ٢	أ٢	,٥٧	,٢٢	,٠٩	(,٩٣)					
	ب٢	,٢٢	,٥٧	,١٠	,٦٨	(,٩٤)				
	جـ٢	,١١	,١١	,٤٦	,٥٩	,٥٨	(,٨٤)			
الطريقة ٣	أ٣	,٥٦	,٢٢	,١١	,٦٧	,٤٢	,٣٣	(,٩٤)		
	ب٣	,٢٣	,٥٨	,١٢	,٤٣	,٦٦	,٣٤	,٦٧	(,٩٢)	
	جـ٣	,١١	,١١	,٤٥	,٣٤	,٣٢	,٥٨	,٥٨	,٦٠	(,٨٥)

ويقترح فسك منذ مطلع السبعينيات أن يقوم الباحثون فى مجال الاختبارات النفسية بتحليل مصفوفات عديدة لمعاملات الارتباط بين مقاييس عدد من التكوينات الفرضية (السمات) لطرق مختلفة، ومن فروضه الأساسية أننا نتوقع فى جميع الحالات أن تكون مقاييس نفس السمة تعطينا نمطا لمعاملات الارتباط يتشابه مع مقاييس سمة أخرى، فمثلا معامل الارتباط بين تقديرات الأقران للسيطرة واستخبار التقرير الذاتى لهذه السمة يظهر على وجه التقريب نفس نمط معاملات الارتباط بين نفس الطريقتين فى قياس سمات المثابرة أو الانتساب أو الاستقلال أو أى سمة أخرى تستخدم فى قياسها هاتان الطريقتان، وقد اختبر هذا الفرض الذى يقترحه فسك فى عدد من البحوث التى أجريت طوال العقدين الماضيين فى مجال مقاييس الشخصية وأظهرت النتائج أن معاملات الارتباط بين السمات فى هذه الحالة لا تظهر إلا تشابها جزئيا مع وجود قدر من الفروق الدالة يرجع إلى اختلاف الطرق (التباين بين الطرق) بل قد نحصل على أنماط مختلفة لمعاملات الارتباط حتى داخل الطرق (التباين داخل الطرق)، وقد وجدت هذه الفروق مثلا بين وسائل التقرير الذاتى التى بنيت مستقلة واختلفت فى القالب (صواب ـ خطأ ـ أو مقارنة ثنائية أو قائمة عبارات) أو فى محتوى المفردات المستخدمة فى الاختبار. ومع ذلك فقد يوجد قدر دال من التباين يمكن قياسه بالطرق المختلفة ويشير إلى تكوين فرضى له معنى سيكولوجى ويمكن تعميمه عبر الطرق المختلفة المستخدمة فى تقويمه نفسيا، وفى هذا وحده معنى صدق التكوين الفرضى.

٣ - **التحليل العاملى:** قد يكون التحليل العاملى من أهم الوسائل التى تستخدمها فى تحديد صدق التكوين الفرضى، وهو فى جوهره أسلوب لتحديد بنية السمات النفسية، وتقوم على تحليل العلاقات بين البيانات الأمبريقية كما تتمثل هذه العلاقات فى صورة معاملات الارتباط، فنقطة البدء فى التحليل العاملى حساب معاملات الارتباط بين درجات كل اختبار والاختبارات الأخرى، التى يجب أن تكون قد طبقت على نفس العينة من المفحوصين، وعندما نفحص الجدول الذى يتضمن هذه المعاملات (والذى يسمى فنيا مصفوفة الارتباط) ونطبق عليه طرقا إحصائية خاصة لايتسع المقام لشرحها قد نلاحظ وجود تجمعات معينة بين الاختبارات مما يوحى بوجود سمات مشتركة بينها، فمثلا إذا كانت الاختبارات التى يرتبط بعضها

ببعض ارتباطا عاليا من نوع اختبارات المفردات والمشابهات والأضداد وتكملة الجمل وكانت معاملات ارتباط هذه الاختبارات بالاختبارات الأخرى منخفضة نستنتج من ذلك وجود عامل إحصائى تتشبع به المجموعة الأولى من الاختبارات تشبعا عاليا وتتشبع به المجموعة الثانية تشبعا منخفضا (والتشبع مفهوم إحصائى يستخدم فى التحليل العاملى ليدل على علاقة المتغير بالعامل).

وبعد ذلك نبدأ عملية تفسير العامل الإحصائى الذى تم الحصول عليه فإذا كانت الاختبارات من نوع الأداء الأقصى كما هو فى مثالنا الحالى يفسر العامل بأنه «القدرة على الفهم اللغوى» وقد يفسر العامل بأنه «مهارة» إذا كانت الاختبارات من النوع الذى يسمى «نفسحركى»، وقد يفسر بأنه «سمة» أو «ميل» أو «اتجاه» أو «قيمة»...إلخ حسب طبيعة الاختبارات التى تنتمى إلى الأداء المميز.

وبالطبع نحن لا نعتمد فى هذا التحليل على مجرد تأمل مصفوفة الارتباط وإلا كانت العملية صعبة ومعقدة وغير مؤكدة بل مستحيلة وخاصة إذا كانت بطارية الاختبارات الكبيرة، ولذلك تستخدم الأساليب الإحصائية الأكثر دقة لتحديد العوامل المشتركة التى يمكن اعتبارها مسئولة عن معاملات الارتباط الناتجة.

وفى التحليل العاملى نلاحظ أن عدد المتغيرات التى نصف فى ضوئها أداء كل فرد تختصر باستمرار من عدد كبير من الاختبارات الأصلية إلى عدد أصغر نسبيا من العوامل أوالسمات المشتركة، ثم يمكن وصف الفرد فى ضوء درجاته فى العوامل القليلة العدد. بعبارة أخرى يمكن القول أن من أهم أغراض التحليل العاملى تبسيط وصف السلوك باختزال عدد المتغيرات أى أنه يسير فى اتجاه العلم الصحيح من التعدد والكثرة التى تتمثل فى متغيرات الاختبارات إلى القلة التى تتمثل فى العوامل المشتركة أو السمات، إلا أن للتحليل العاملى وظيفة أخرى هى فى رأينا أكثر ارتباطا بمفهوم صدق التكوين الفرضى وهى محاولة اختبار صحة فروض معينة حول البنية العاملية للاختبار الواحد من خلال تحليل مصفوفة ارتباط مفرداته أو أسئلة أو بطارية من الاختبارات النفسية، وهى وظيفة أشار إليها فؤاد أبو حطب (١) منذ مطلع السبعينيات ثم ازداد الاهتمام بها فى الوقت الحاضر فيما يسمى التحليل العاملى المدعم Confirmatory Factor Analysis.

وبعد تحديد العوامل يمكن أن نفيد منها فى وصف التكوين العاملى للاختبار. فكل اختبار يمكن أن يوصف فى ضوء العوامل الأساسية التى تحدد درجاته بالإضافة إلى تشبعه بكل عامل، وتشبعات العوامل هذه تمثل بدورها معاملات ارتباط بين الاختبار والمعامل، وهذا المعامل يسمى الصدق العاملى الذى هو فى الواقع معامل ارتباط بين الاختبار وبين ما هو شائع أو مشترك أو عام فى مجموعة الاختبارات أو غيرها من المؤشرات الدالة على السلوك، ومن المؤكد أن المتغيرات التى نقوم بتحليلها عامليا تتضمن معلومات نشتقها من الاختبارات أو غيرها من أدوات التقويم النفسى كالتقديرات والأحكام، وهذه جميعا يمكن أن تستخدم فى محاولة استكشاف الصدق العاملى لاختبار معين وتحديد السمات المشتركة التى يقيسها ، كما يمكن أن يطبق التحليل العاملى على مصفوفات ارتباط المفردات التى يتألف منها الاختبار الواحد.

٥ - **الإجراءات التجريبية**: من الوسائل الأخرى التى يمكن أن نلجأ إليها لتحديد صدق التكوين الفرضى ما توفره لنا الإجراءات التجريبية من معلومات حول أثر متغيرات معينة على درجات الإختبار. فإذا كان الإختبار التجريبى (الجديد) معدا لقياس القابلية للقلق مثلا يمكن أن نطبق هذا الاختبار على مجموعة من الأفراد فى موقف يثير القلق (مثل الامتحان فى ظروف ضاغطة) ثم يحسب معامل الارتباط بين درجات الأفراد فى اختبار القلق قبل التجريب وبين مختلف المؤشرات الفسيولوجية وغيرها التى تدل على القلق خلال التجريب (موقف الامتحان) وبعده، ويمكن القول أن هذا الإجراء التجريبى يشبه إلى حد كبير ما نلجأ إليه فى تقدير الصدق التنبؤى، باستثناء الدور الإيجابى للباحث فى تقديم مثير القلق والتحكم فيه، أما المقاييس المباشرة الدالة على القلق فيمكن اعتبارها محكا نراجع عليه تنبؤاتنا من درجات الاختبار.

وقد تفيد الوسائل التجريبية أيضاً فى تحقيق الفروض المختلفة حول الاختبارات، ففى مثالنا السابق ذكره يمكن تطبيق اختبار القلق قبل التجريب (إعطاء خبرة القلق) وبعده تم نتحقق مما إذا كانت درجات الإختبار ترتفع ارتفاعا له دلالة عند إعادة الاختبار وبالطبع يمكن استخدام تصميمات تجريبية تعتمد على المجموعات الضابطة والتجريبية، فإذا كانت النتيجة إيجابية فإن ذلك يدل على أن درجات الاختبار تعكس المستوى الراهن للقلق، وتبرهن على صدق التكوين الفرضى الذى يقيسه.

جدول رقم (١٤) يوضح أنواع الصدق واستعمالات كل منها

نوع الصدق	السؤال المطروح	الإجـــراء	الاستخدام الرئيسى	أمثلة
صدق المحتوى	هل اكتسب المفحوص مقدارا كافيا من المعلومات أو المهارة من البرنامج التدريبى الذى تعرض له؟	نقارن بين أسئلة الاختبار من الوجهة المنطقية بالمحتوى الذى نفترض أن أسئلة الاختبار تنتمى إليه.	اختبارات التحصيل المدرسى والكفاءة المهنية.	فحص اختبار القدرة على الاختزال لمعرفة هل يحتوى مضمونه بالفعل على مهارات العمل المكتبى.
الصدق المرتبط بالمحكات (١) الصدق التنبؤى	هل تتنبأ درجات الاختبار بأداء معين له أهمية فى المستقبل؟	إعطاء الاختبار ثم نحصل على مقياس لمحك بعد فترة زمنية ثم نقارن الاختبار بالمحك.	الاختبارات التى تستخدم فى الانتقاء والتصنيف.	مقارنة اختبارات القبول بالكليات أو المعاهد بتحصيل الطلاب فيما بعد.
الصدق المرتبط بالمحكات (٢) الصدق التلازمى	هل تشخص درجات الاختبار صعوبات الأداء فى الوقت الحاضر؟	نعطى الاختبار ثم نحصل على مقياس مباشر دون فاصل زمنى ثم نقارن الاختبارات بالمحك.	الاختبارات التى تستخدم فى أغراض التشخيص التربوى أو الكلينيكى.	مقارنة أداء المتخلفين عقليا بالأسوياء فى اختبار للذكاء.
صدق التكوين الفرضى	ما المعنى النفسى للمفهوم الذى يقيسه الاختبار؟	صياغة الفروض حول الاختبار ثم التحقق منها بالإجراء التجريبى أو الإحصائى المناسب.	الاختبارات التى تستخدم فى أغراض البحث العلمى حول الاختبارات النفسية.	دراسة اختبار القدرة الموسيقية لتحديد إلى أى حد تتأثر الدرجات فيه بعوامل التدريب الموسيقى.

العوامل المؤثرة فى الصدق:

إذا صيغ الصدق فى صورة معامل ارتباط فإنه يصبح مثل معامل الثبات يتأثر بالعوامل المختلفة التى تؤثر فى معامل الارتباط وخاصة ما يتصل بتجانس العينات والتى تناولناها بالتفصيل فى الفصل السابق إلا أن هناك عوامل تؤثر فى الصدق تأثيرا خاصاً ولابد من أن توضع فى الاعتبار.

١ - **طبيعة عينة التقنين:** من المهم لفهم طبيعة معامل الصدق وصف طبيعة عينة التقنين وصفا مفصلا، لأن الاختبار الواحد قد يقيس عمليات نفسية مختلفة إذا طبق على عينات تختلف فى العمر أو الجنس أو المهنة أو المستوى التعليمى أو أى متغير آخر يرتبط بمفهوم الصدق. وعلى ذلك فإن الاختبار الواحد قد يكون له صدق مرتفع فى التنبؤ بمحك معين فى بعض العينات، بينما ينخفض صدقه بالنسبة إلى عينات أخرى وربما يكون مقياسا صادقا لعمليات نفسية مختلفة فى عينتين مختلفتين، ومن ذلك مثلا اختبار العمليات الحسابية الذى قد يكون مقياسا صادقا للاستدلال الحسابى عند أطفال الصفوف الدنيا فى المرحلة الابتدائية، بينما يقيس السرعة العددية بصدق عند طلاب المرحلة الثانوية.

٢ - **طبيعة العلاقة بين الاختبار والمحك:** من المعروف أن حساب معامل الارتباط بين الاختبار والمحك، إذا استخدمت فيه طريقة كارل بيرسون أو تقريباتها المختلفة، فإن ذلك يتضمن افتراضا جوهريا وهو أن العلاقة بين المتغيرين خطية linear على امتداد مدى الفروق الفردية فيهما. وبالطبع فإن الافتراض قد يكون صحيحا فى عدد كبير من الحالات، إلا أنه فى بعض الحالات قد لا تكون العلاقة كذلك، ومن ذلك مثلا حين يتطلب أحد الأعمال حدا أدنى من فهم القراءة بعين العاملين فيه على قراءة بعض التعليمات وفى هذه الحالة نجد أنه حالما توافر هذا الحد الأدنى فإن أى زيادة عليه فى القدرة على فهم القراءة يصبح غير مرتبط بدرجة النجاح فى العمل، وتوصف العلاقة بين الاختبار والمحك فى هذه الحالة بأنها غير خطية non-linear ومن أمثلة ذلك العلاقة المنحنية curvilinear والمثلثة triangle أوالعلاقة التى تدل على اختلاف التباين heteroscedasticity، ومن أمثلة النوع الأخير من العلاقات حين نجد أن الأداء فى اختبار الذكاء يكون شرطا ضروريا ولكنه

ليس كافيا للنجاح المدرسى، وفى هذه الحالة نجد التلاميذ ذوى الذكاء المنخفض يؤدون أداء سيئا فى المدرسة، ولكننا حين نفحص أداء التلاميذ من ذوى الذكاء المرتفع سوف نجد أن بعضهم يؤدون أداء جيدا فى المدرسة وبعضهم الآخر قد يؤدى أداء سيئا لأسباب تتصل بنقص الدافعية، أوفقدان الميل أو غيرها، وفى هذه الحالةنجد تباينا فى التحصيل بين ذوى الدرجات المرتفعة فى الذكاء أكثر اتساعا مما نجده بين ذوى الدرجات المنخفضة فيه.

٣ - **أثر ثبات الاختبار فى صدقه:** من بديهيات القياس النفسى أن الاختبار غير الثابت وغير الدقيق لايمكن أن يكون صادقا. لأن معنى انخفاض ثبات الاختبار أن تباين الخطأ فيه أكبر من التباين الحقيقى، وبالطبع فإن تباين الخطأ فى درجات المفحوصين لايمكن أن يرتبط ارتباطا دالا (أى يتجاوز مستوى الصدفة) مع أى محك، وعلى هذا كلما زاد تباين الخطأ يقل معامل الصدق، لأن الصدق هو التباين الحقيقى للاختبار أما تحرر الاختبار تماما من تباين الخطأ فإنه يصل بالثبات إلى حد التمام ويصبح التباين الكلى فى هذه الحالة مساويا للتباين الحقيقى، وفى هذه الحالة النظرية المستحيلة الحدوث وحدها يمكن للثبات أن يساوى الصدق.

إلا أن الاختبار النفسى - كغيره من صور القياس - يكون فى العادة مشوبا بقدر من تباين الخطأ، ولهذا فإن من المتوقع فى هذه الحالة الا يزيد معامل صدقه عن معامل ثباته، لأنه لايكون منطقيا على أى نحو أن يرتبط الاختبار مع نفسه أو مع ما يقابله أو يشبهه من صور متكافئة (وهو معنى الثبات) ارتباطا أقل من ارتباطه مع اختبارات أو مقاييس أخرى مختلفة عنه هى ما يسمى المحكت (وهو معنى الصدق)، ولذلك فكثيرا ما يقال أن الحدود القصوى لمعاملات الصدق المحسوبة هى معاملات الثبات والتى عادة ما لا تتجاوزها النتائج الإمبريقية للاختبارات.

إلا أن هذه العلاقة بين الثبات والصدق قد أسىء عرضها فى عدد من البحوث النفسية التى أجريت فى مصر والعالم العربى فى السنوات الأخيرة، وخلاصة هذا الخطأ أن بعض الكتب الإحصائية باللغة العربية أشارت إلى مفهوم مؤشر الثبات in-dex of reliability وهو عبارة عن الجذر التربيعى لمعامل الثبات بمصطلح الصدق الذاتى، وهى تسمية غير دقيقة فى رأى كاتب هذا الفصل(*)، لأنها تحمل المؤشر ما

(*) أ.د. فؤاد أبو حطب.

لايعليه، فهو لاينتمى إلى الصدق بمعناه العلمى الدقيق.

وتلقف باحثونا المستعجلون الفكرة ووجدوا فيها تعويضاً عن عجزهم عن البحث فى الطرق الصحيحة لتحديد صدق الاختبارات، فنجد الواحد منهم يحسب ثبات اختباره ثم يحسب الجذر التربيعى لهذا الثبات ويشير إليه بالصدق، وكفى الله المؤمنين شر القتال. لون آخر من ألوان العبث العلمى الذى يشهده المسرح المعاصر لعلم النفس.

وخلاصة فكرة مفهوم مؤشر الثبات (أو ما يشار إليه خطأ بالصدق الذاتى) أن مربع معامل الصدق (أى مربع معامل الارتباط بين الاختبار والمحك) يجب أن يكون أقل من معامل الثبات، وذلك لأن معامل الثبات يساوى مربع معامل الارتباط بين الاختبار ومتوسط الأصل الإحصائى العام الذى تشتق منه العينة، ومعنى ذلك أن مربع معامل الصدق يجب أن يكون أقل من مربع معامل ارتباط الاختبار بمتوسط الأصل الإحصائى، وبعبارة أخرى أن الأختبار لايمكن أن يرتبط بمتغير خارجى (وهو المحك) أعلى من ارتباطه بمتوسط أصله الإحصائى العام.

لتوضيح هذه الفكرة نعطى المثال التالى: نفرض أن معامل ثبات أحد اختبارات الذكاء بلغ ٠,٨٧ ، معنى ذلك أننا لو استبعدنا أخطاء العينة، فإن هذا الاختبار لايمكن أن يتنبأ بالتحصيل المدرسى (وهو المحك) بمربع معامل صدق يزيد عن هذا المقدار (أى ٠,٨٧)، وعلى ذلك فإن معامل الصدق لايجب أن يزيد بحال من الأحوال عن الجذر التربيعى لمعامل الثبات (وهو فى هذه الحالة ٠,٩٢) ان هذا الجذر التربيعى ليس معامل صدق) وإنما يدل على الحد الذى لايمكن أن تتجاوزه معاملات الصدق، أنه قد يعنى أن معامل الصدق قد يكون أعلى من معامل الثبات - فى ضوء هذا الحد الأقصى - إلا أن هذا الافتراض النظرى لم يتحقق تجريبيا أبدا، وجميع البيانات الامبريقية التى يجمعها الباحثون بنزاهة تتفق مع البديهية التى دارت حولها مناقشتنا السابقة وهى أن معاملات الصدق لابد أن تكون أقل من معاملات الثبات.

الفصل السادس

تصحيح الاختبارات وتفسير الدرجات

تصحيح الاختبار:

يعد تصحيح الاختبارات بإعطاء الفرد درجة أو تقديرا، وتفسير هذه الدرجة هو خطوة مهمة، ولو أنه فى ذاته يعد مقدمة لاتخاذ قرار عملى أو تفسير علمى عن الفرد أو مجموعة الأفراد موضوع القياس. وعلى كل فإن مشكلات تصحيح الاختبارات لا تقل أهمية وخطورة عن المشكلات الأخرى التى أثرناها فى هذا الباب ولعل أخطر ما يرتبط بهذه المشكلة ما يتصل بموضوعية التصحيح والموضوعية objectivity هى مشكلة العلم عامة، ومشكلة العلوم الإنسانية خاصة، وقد تناولها فى السنوات الأخيرة عدد من الفلاسفة وعلماء النفس (٣، ٤، ٥، ٦)، ودون أن ندخل فى تفاصيل مناقشة شائكة حول هذه المسألة، نشير هنا إلى أننا نتناول الموضوعية فى إطار محدد وهو تصحيح الاختبارات، صحيح أن المسألة تتداخل مع جميع مراحل بناء الاختبار وتطبيقه وتقنينه، وكلها تتصل بدور الباحث أو الفاحص أو المصحح.

وبالطبع فإن المعنى المباشر للموضوعية هو تحديدها «سلبيا» ولهذا كثيرا ما نجدها تعرف بتحرر الفاحص من العوامل الذاتية كالتحيز والتعصب والآراء المسبقة والأفكار الجاهزة التى يحاول فرضها على الوقائع، إلا أن هذا التحديد «السلبى» ليس تعريفا مفيدا للمفهوم، وكل ما يعطيه لنا من معنى هو أن الموضوعية مقابل الذاتية، وفى رأى فؤاد أبو حطب أن هذه المقابلة الثنائية ليست هى التناول الصحيح للمسألة، فلغة الأضداد ليست هى اللغة الصحيحة للعلم، ولا لمنهجه والأصح أن نتصورها على هيئة متصل يمتد على خط مستقيم يبدأ بالموضوعية الكاملة (شبه الآلية الميكانيكية) وينتهى بالذاتية الكاملة (الاعتماد الكامل على الهوى والغرض) وبينها توجد درجات كالسمية عند سيد عثمان (٤) «البصيرة العلمية» التى تتأتى بالخبرة.

وعلى وجه العموم نحن فى حاجة فى مجال الاختبارات النفسية أن تكون عمليات تطبيق الاختبار وتصحيحه وتفسير درجاته مستقلة نسبيا عن الحكم الشخصى للفاحص، حتى ولو توافر له قدر كاف من هذه البصيرة العلمية، ولكى نتأكد من ذلك فإننا عادة ما نلجأ إلى تحديد درجة «اتفاق الملاحظات والأحكام اتفاقا مستقلا» وهذا

هو التعريف الذى اقترحه فؤاد ابو حطب للموضوعية، وهو تعريف موجب يتجاوز التعريفات «السالبة» المعتادة، بالإضافة إلى أنه قابل للتناول والمعالجة، فهذا «الاتفاق المستقل» فى الملاحظات والأحكام والتقديرات يمكن تحديده بحساب معامل الإرتباط بين عدد من الفاحصين فى ملاحظتهم أو حكمهم أوتقويمهم - مستقلين بعضهم عن بعض - لنفس المفحوصين، وفى هذه الحالة تصبح «الأهواء الذاتية «للفاحصين جزءا من تباين الخطأ، ويدخل الأمر كله فى قضية ثبات الاختبار وهو نوع من الثبات وثيق الصلة بموضوعنا الراهن.

كما يمكن تحديد موضوعية الفاحص الواحد «بطريقة مماثلة وذلك بحساب معامل الارتباط بين عدد من ملاحظاته أو أحكامه أو تقديراته أو تقويماته لنفس المفحوصين والتى تختار كعينات من سلوكه، ومن المؤكد أن هذا المعامل إذا كان منخفضا فإنه يدل على ذاتية المصحح» أما إذا كان مرتفعا فقد يكون مؤشرا على «البصيرة العلمية» لديه، بشرط أن نتأكد من عدم تداخل ملاحظاته بعضها مع بعض، وإلا كان معامل الارتباط المرتفع هذا مؤشرا زائفا على هذه البصيرة، فقد يدل فى هذه الحالة على اتساق التحيزات أو الأهواء الخاصة وفى جميع الحالات يجب التنبيه إلى أن الموضوعية مسألة درجة، وإن كانت أكثر وضوحا فى نمط الاختبارات الذى نسميه «امتحانات المقال» ففى هذا النوع كثيرا ما يتساءل الشخص لماذا حصل على درجة منخفضة، ومن المؤكد أن السبب الرئيسى هو أنه فى هذا النوع من الامتحانات يكون من الصعب تحيد الإجابة الجيدة وبالتالى يكون من الصعب إعطاء درجة مناسبة لاستجابة ليست صحيحة صحة كاملة.

وبالطبع لايمكن للباحث النفسى أن يجرى دراسة للسلوك كما لايمكن أن يتخذ قرارا عمليا إذا لم تكن معايير التصحيح على درجة من الموضوعية ولذلك يقترح البعض لحل هذه المشكلة اما تحديد قواعد للحكم يلتزم بها المصححون للاختبارات التى تتطلب إجابات حرة أو استخدام الأسئلة الموضوعية التى تتطلب التعرف على الإجابة الصحيحة من بين عدة إجابات، ويستخدم فى تصحيحها مفاتيح معدة مقدما.

تصحيح الأسئلة التى تتطلب استجابات حرة:

لايزال القياس الفردى يستخدم أسئلة ومشكلات تتطلب درجة من الحكم أثناء

التصحيح، ومع ذلك فمن الممكن إعداد بعض الوسائل التى تتوفر فيها درجة من الموضوعية فى هذا التصحيح وبخاصة فيما يتعلق بالخصائص الرئيسية للسلوك، فعند تقويم سلوك معين مثل الخط، أمكن للباحثين إعداد نموذج للتصحيح يحتوى على عينات من الخطوط تمثل مستويات مختلفة من الجودة وعلى المدرس أن يحدد العينة الأكثر شبها بخط التلميذ حتى يحدد الدرجة التى يعطيها له وكذلك يمكن إعداد مقاييس معتمدة على تقدير الإنتاج للحكم على «نوع» بعض ألوان السلوك الصناعى أو التجارى أو الإدارى، ومع أن تحديد قواعد تصحيح بعض الاختبارات التى تحتوى على استجابات حرة حقق النجاح فى اختبارات المهارات العملية، إلا أن هذه الوسائل لم تلجح نجاحا مؤكدا فى تصحيح الاختبارات اللغوية، ومع ذلك فإن التباين بين المصححين يمكن تقليلة إذا أعددنا بروتوكولات توضح نمط التصحيح المتفق عليه بالنسبة للإجابات النموذجية، ومن أمثلة ذلك نماذج التصحيح الخاص باختبار ستانفورد - بينيه للذكاء، وبروتوكولات اختبار رورشاخ واختبار تفهم الموضوع، وقواعد التصحيح فى اختبارات الابتكار.

وعلى الرغم من أن بروتوكولات التصحيح قد تكون مناسبة لأغراض القياس، إلا أننا نؤكد بعض الحذر حول تصحيح اختبارات الاستجابات الحرة بوجه عام، فمعظم مصادر تباين الخطأ فيها تنشأ مما يمكن تسميته «المعادلة الشخصية» للمختبرين أوالمصححين، فقد لوحظ مثلا أن بعضهم يعطى درجات مرتفعة بصفة مستمرة بينما يميل البعض الآخر إلى إعطاء درجات منخفضة، كما أن الدرجة التى يعطيها المختبر أو المصحح تتأثر باستعداده العقلى وتحيزه وتخمينه وأغراضه الشخصية وأهوائه، وفى هذه الحالة لامناص من أن يقوم بالتصحيح أكثر من مصحح واحد تتوافر فيهم جميعا خصائص التدريب الجيد على بروتوكولات تصحيح هذه الاختبارات أو تقدم لهم قواعد التصحيح واضحة بحيث يمكنهم تطبيقها دون الوقوع فى أخطاء سوء التأويل.

ولعل من المصادر الأخرى لتباين الخطأ فى هذا النوع من الاختبارات ما يسمى أثر أو وضع الهالة الذى يحدثه المفحوص ويمكن التقليل من هذا الأثر إذا لجأنا إلى ما يسمى التصحيح «الأعمى»، أى يقوم المختبر أوالمصحح بتقدير الدرجات للأفراد دون أن يعلم نتائج الاختبارات السابقة التى أجريت عليهم أو بالفئة التى ينتمى إليها

المفحوص فى البحث (المجموعة التجريبية أو الضابطة) أوالمستوى الذى هو عليه (العقوق أو التخلف) وبدون هذا الاحتياط يصعب علينا أن نتفادى تأثير هذا العامل الذاتى فى الدرجة المعطاة.

تصحيح الأسئلة التى تتطلب الاختيار من متعدد:

يتكون نموذج التصحيح لما نسميه اختبارات التعرف من قائمة الإجابات الصحيحة (أو مفتاح التصحيح) يمكن أن يستخدمه الفاحص بل نكاد نقول أن أى شخص يمكنه أن يستخدمه فى تصحيح هذه الاختبارات، وقد اخترعت عدة وسائل أكثر كفاءة فى تسجيل الاستجابات فى هذا النوع من الاختبارات وتصحيحها، ومنها النسخ الكربونية، وفيها نجد أن الوجه الخلفى لصفحة الأسئلة (أو ورقة الإجابة) تلتصق به وتخفيه ورقة كربونية، ويكون مفتاح الإجابة مطبوعا على هذا الوجه الخلفى ويختار المفحوص الإجابة ثم يضع العلامة التى تدل على هذا الوجه الظاهر لورقة الأسئلة (أو الإجابة) وبالطبع فإن هذه العلامة تطبعها ورقة الكربون على الوجه الخلفى للورقة، وعند التصحيح يقوم المختبر أو المصحح بفض الوجه الخلفى الذى توجد فيه مربعات مطبوعة توضح أين يجب أن تكون الإجابات الصحيحة (أى مفتاح الإجابة)، وما على المصحح فى هذه الحالة إلا عد العلامات الكربونية التى طبعت فى المربعات الصحيحة، وبالطبع فإن هذا الإجراء مكلف اقتصاديا.

ومن الإجراءات الشائعة استخدام أوراق مستقلة للإجابة، وهو إجراء يؤدى إلى اختصار التكاليف، لأن كتيب الاختبار يمكن استخدامه فى هذه الحالة أكثر من مرة، كما أن الإجابات يمكن تصحيحها بسهولة بمفتاح مثقوب أو بواسطة الآلات الحديثة.

وقد ازداد فى الوقت الحاضر الاعتماد على تصحيح الاختبارات تصحيحا أوتوماتيكيا إلى حد أن برامج كاملة للتقويم النفسى فى أوربا والولايات المتحدة تتم بالحاسبات الالكترونية، وتختلف طرق التصحيح الآلى إلا أنها تتشابه فى عدد من العناصر منها ضرورة وجود ورقة إجابة بمواصفات معينة، وأن يقوم المفحوص بتسويد مكان الإجابة بالقلم الرصاص، وتقوم الآلة بقراءة الإجابة بواسطة نظام للشفرة الموضعية بإحدى طريقتين أولاهما خلايا ضوئية حساسة للأسود والأبيض، وثانيهما التوصيل الكهربائى لمادة الجرافيت (التى يصنع منها الأقلام)، وفى الطريقة الأخيرة

نجد الآلة تحتوى على «أصابع» مكهربة توضح علامات القلم الرصاص كما تحتوى الآلة على عداد تسجيل العدد الكلى للعلامات الموضوعية فى أماكنها الصحيحة، كما أن الآلة يمكنها أن تسجل عدد الأخطاء، كما يمكنها أن تصحح الدرجة من أثر التخمين، ويمكن لبعض هذه الآلات أن يصحح عدة آلاف من أوراق الإجابة تصحيحا دقيقا فى الساعة الواحدة. والحاسب الالكترونى يمكنه أن يقرأ ورقة الإجابة مرة واحدة ثم يقارن إجابات المفحوص بالمفتاح المختزن فى ذاكرة الكومبيوتر ،.وتطبع درجة المفحوص إما على ورقة الإجابة ذاتها أو فى بطاقة خاصة، كما يمكن للحاسب الالكترونى أن يقدم درجات المفحوصين مبوبة (مصنفة على النحو الذى يرغبه الفاحص). كما يمكنه أن يقوم بتحليل هذه الدرجات مباشرة بالطرق الإحصائية الملائمة.

والصعوبة الرئيسية التى تواجهها آلات تصحيح الاختبارات، والتى يهتم بها المختصون فى تطبيق الاختبارات، هى أن الآلات لايمكنها أن تصحح تصحيحا دقيقا ما لم تكن الأماكن المخصصة للإجابة قد قام المفحوص بتسويدها تسويدا كاملا ونظيفاً وحسب المواصفات المطلوبة وألا يكون قد وضع إجاباته فى غير مواضعها، فمن المعروف أن الآلة تعطى وحدة الدرجة عندما تظهر العلامة التى وضعها المفحوص فى المكان الصحيح ومع ذلك فنحن نتوقع أن تكون العلامة التى يضعها المفحوص بالقلم الخفيف أوالثقيل، أو تكون عريضة أو ضيقة، ولذلك فلابد أن يضع المفحوص هذه العلامة بقلم معين وأن يملأ المساحة الكاملة المخصصة للإجابة فى ورقة الإجابة حتى نتأكد أن الآلة سوف تعدها ولا تتجاهلها، ومن ناحية أخرى فإن الآلة قد لا تتجاهل العلامات غير الكاملة والمحو غير النظيف للأخطاء، وقد يكون عدد الإجابات فى ورقة الإجابة التى لا تتوفر فيها المواصفات اللازمة كثيرة، ولذلك نجد فى بعض المؤسسات يحددون موظفين يقومون بفحص كل ورقة قبل أن تسلم للآلة لتصحيحها، وهؤلاء يقومون بمراجعة اجابات المفحوصين لا بفرض تصحيحها ولكن بغرض التأكد من أن طريقة الإجابة مناسبة للتصحيح، ومعنى ذلك زيادة تكاليف التصحيح، ولاشك أنه يمكن التغلب على هذه الصعوبة بإجراء الاختبار وتطبيقه على نحو سليم.

تصحيح أثر التخمين:

من المشكلات المهمة فى تصحيح الاختبارات التى تتطلب اختيار الإجابة من

بديلين أو أكثر ما يتصل بالتخمين guessing، وهو مصطلح عام يشير إلى أنماط السلوك التى تصدر عن المفحوص حين يستجيب لأحد البدائل الاختيارية للإجابة على السؤال الذى لا يعرف إجابته الصحيحة، وهذه الأنماط السلوكية عديدة، تشمل من بين ما تشمل ما يلى (٣٤ : ٥٩).

١ – حذف بديل أو أكثر من بين البدائل الاختيارية للإجابة على السؤال والحكم عليها بأنها خاطئة تماما، أو حين يكون المفحوص غير متأكد من صوابها.

٢ – الاستفادة من بعض المنبهات غيرالمقصودة والمتاحة بسبب صياغة السؤال أو البدائل الاختيارية، وقد تكون هذه المنبهات فى المعنى أو فى البنية النحوية.

٣ – وقوع المفحوص فى الشرك الذى ينصبه له مؤلف الاختبار حين يطرح بديلا جذابا ولكنه خاطئ.

٤ – الاستجابة على أساس ما يوجد فى أحد البدائل الاختيارية من عناصر تجذبه دون أن يكون لديه ثقة كبيرة فى صحته.

٥ – استخدام طريقة عشوائية فى الاستجابة على نحو يعتمد كلية على المصادفة، وقد يصل ذلك فى تطرفه إلى حد انتقاء البدائل الاختيارية التى تقع فى الاختبار فى موضع معين: مثل انتقاء الإجابة (أ) أو (جـ) أو (هـ) فى جميع الأسئلة.

ومن المؤكد أن المفحوص إذا لجأ فى إجابته إلى محض التخمين معتمدا فى ذلك على العشوائية الكاملة فإنه قد يحصل على بعض الدرجات، لأن بعض تخميناته قد تقع بالمصادفة على الإجابات الصحيحة، ولعل هذه المسألة تزداد وضوحا إذا علمنا أنه فى بناء الاختبار الموضوعى الجيد الذى تقوم أسئلته على الاختيار من متعدد عادة ما تتوزع الإجابات الصحيحة على البدائل المختلفة بالتساوى، فمثلا إذا كانت أسئلة الاختبار من نوع البديلين فإن ٥٠٪ من الإجابات الصحيحة يكون رمزها (أ) مثلا، ٥٠٪ من هذه الإجابات يكون رمزها (ب)، على أن تتوزع عشوائيا، دون أى تتابع منظم فى الاختبار كله،و فإذا كانت أسئلة الاختبار من نوع البدائل الثلاثة فإن ٣٣٪ من الأسئلة يكون إجاباتها الصحيحة (أ)، ٣٣٪ رمزها (ب) ، ٣٣٪ رمزها (جـ) على أن تتوزع مرة أخرى عشوائيا فى الاختبار، وبالمثل إذا كانت البدائل أربعة تتوزع

الإجابات الصحيحة لكل (٢٥ ٪ من الأسئلة على أحد الرموز (أ)، (ب)، (جـ)، (د)، وهكذا.

لو افترضنا أن مفحوصًا اعتمد على العشوائية الكاملة فاختار الإجابات التى رمزها (أ) فى جميع الأسئلة، إنه فى هذه الحالة يحصل دون حق على الدرجة ٥٠ فى الحالة الأول، ٣٣ فى الحالة الثانية، ٢٥ فى الحالة الثالثة.

وحيث أننا لانعرف فى هذه الحالة عدد التخمينات الصحيحة التى توصل إليها المفحوص بهذه الطريقة، فإننا نستطيع استنتاج هذا العدد من عدد التخمينات الخاطئة والتى تمثلها المعادلة الآتية:

$$\text{خ خ} = \frac{\text{خ}}{\text{ن} - ١}$$

حيث خ خ = عدد التخمينات الخاطئة.

خ = عدد الإجابات الخاطئة فى الاختبار كله.

ن = عدد البدائل الاختيارية.

ويمكن استخدام المعادلة السابقة فى معادلة أشمل لتصحيح أثر التخمين والتى تستخدم فى تقدير عدد الأسئلة التى يعرف المفحوص إجاباتها الصحيحة بالفعل ويستحق عليها الدرجة الاختبارية وهذه المعادلة هى:

$$\text{د} = \text{ص} - \left(\frac{\text{خ}}{\text{ن} - ١}\right)$$

حيث :

ص = عدد الإجابات الصحيحة فى الاختبار كله

د = الدرجة المصححة من أثر التخمين.

خ = عدد الإجابات الخاطئة فى الاختبار كله

ن = عدد البدائل الاختيارية

تفسير الدرجات

الدرجات الخام:

تعطينا معظم الاختبارات وصفا كميا مباشراً لأداء الشخص، ويطلق على الوصف الكمى للأداء اسم الدرجة الخام raw score وقد تكون الدرجة الخام عدد الأسئلة التى أجاب عليها المفحوص إجابة صحيحة أوخاطئة أوالزمن الذى استغرقه فى الأداء أو النسبة المئوية لعدد الإجابات الصحيحة أوالخاطئة على العدد الكلى لأسئلة الاختبار. ويفسر معظم الأفراد الدرجة الخام ويستخدمونها دون معرفة لجوانب القصور فيها لأن هذه الدرجات تتميز بالسهولة فى الحصول عليها وبألفتنا لها من خبرتنا الطويلة بالامتحانات المدرسية، ومع ذلك فإننا نستطيع القول أن الدرجة الخام فى الاختبار النفسى لامعنى لها فى ذاتها وليست لها أية دلالة، ولايمكن أن تفسر إلا باللجوء إلى أساس آخر للمقارنة، ولعلنا نعود بهذا إلى التعريف الأساسى للاختبار النفسى. ومن المهم أن نؤكد فى هذا الصدد أننا لانستطيع تفسير الدرجات فى الاختبار النفسى كما نفعل بالقيم التى نحصل عليها من مقياس الظواهر الطبيعية (الطول مثلا)، فمقياس الظواهر الطبيعية هى من نوع مقاييس النسبة، أى لها صفر مطلق وتتألف من وحدات متساوية، ولذلك فإن هذه المقاييس تسمح لنا بتفسير مباشر للقيم التى نحصل عليها منها، فنقول مثلا أن طول الطفل (س) نصف طول والده (ص) باستخدام درجات المتر مباشرة، أما فى القياس النفسى فإننا لا نستطيع مثل هذا القول لأن المقاييس فى هذه الحالة لا تكون فى أحسن حالاتها إلا من نوع مقاييس المسافة، لنفرض مثلا أن الطفل (س) حصل على درجة مقدارها ١٠ ٪ من أسئلة اختبار، فى معانى الكلمات فهل يعنى ذلك أن ذلك الطفل لايعرف إلا عشر الكلمات التى يجب أن يعرفها؟ الإجابة على ذلك بالنفى لأن الفاحص قد لا يكون مهتما إلا بالكلمات الصعبة غير المألوفة للطفل ولهذا نجد أن محصوله اللفظى قد يكون أكبر فى اختبار أقل صعوبة، وحتى لو حصل الطفل على الدرجة صفر فى هذا الاختبار فإن ذلك يعنى أن الطفل ليست لديه قرة لغوية على الإطلاق، فالفرق بين (س) الذى حصل على الدرجة صفر فى الاختبار والطفل (ص) الذى يحصل على النهاية القصوى أى يجيب على أسئلة الاختبار جميعا، هو فرق فى القدرة على معرفة عينة من الكلمات يتكون منها الاختبار اشتقت من بين آلاف الكلمات، فالدرجة التى تدل على ٦٥ ٪ صواب فى

أحد الاختبارات قد تتساوى مع الدرجة التى تساوى ٣٠٪ فى اختبار آخر أصعب ، ومع ٨٠٪ فى اختبار ثالث أسهل، فمستوى صعوبة الكلمات التى يتألف منها الاختبار تحدد الدرجة الخام للمفحوص.

معنى ذلك أن الفرق بين الأفراد فى الدرجات الخام للاختبارات النفسية لايدل بالضرورة على وجود مسافات «حقيقية» بين الأفراد لأن هذه الاختبارات لايتحقق فيها الشرطان الأساسيان لمقاييس المسافة وهما الصفر المطلق والوحدات المتساوية . فمثلا فى اختبار الذكاء قد يحصل التلميذ (أ) على الدرجة ٥٣ والتلميذ (ب) على الدرجة ٥٦ والتلميذ (جـ) على الدرجة ٥٩، أى تكون الفروق بين الدرجات الخام متساوية، ومع ذلك لايدل هذا على أن الفرق بين (جـ) و(ب) يساوى تماما الفرق بين (ب) و(أ)، ولايمكن الإجابة على مثل هذا السؤال إجابة مؤكدة، لأن الفروق بين الدرجات إما تعتمد على طبيعة الأسئلة (أو الوحدات) التى تكون الاختبار كما قلنا، ولهذا لابد من أن نلجأ إلى بعض الأسس التى تتجاوز الدرجات الخام فى المقارنة بين المفحوصين وتفسير درجاتهم والحكم على أدائهم.

وتضيف آمال صادق وفؤاد أبو حطب (١) هذه الأسس التى يجب استخدامها فى المقارنة ثم التفسير والحكم على الدرجة الخام إلى ثلاثة أنواع هى: المعايير norms، والمحكات Criteria ، والمستويات Standards وفيما يلى تفصيل الأسس الثلاثة:

(أولا) المعايير

كانت «المعايير» أول ما ظهر فى ميدان القياس النفسى للتغلب على الصعوبات المتضمنة فى تفسير الدرجات الخام والحكم على أداء المفحوص وتعد المعايير أساسا لتفسير أداء المفحوصين والمقارنة بينهم فى ضوء أداءهم الفعلى، وتتحدد فى ضوء الخصائص الواقعية لهذا الأداء، وتعتمد فى جوهرها على الأداء الاختبارى لعينة ممثلة للأصل الإحصائى موضع الاهتمام تسمى عينة التقنين Standardization sample وهكذا تتحدد المعايير تجريبيا بما تستطيع مجموعة ممثلة من الأفراد أداءه (عينة ممثلة)، ثم نشير إلى الدرجة الخام التى يحصل عليها المفحوص فى ضوء توزيع الدرجات التى تحصل عليها هذه العينة ليتحدد موقعه ومكانته فى هذا التوزيع: هل

تتفق درجته مثلا مع الأداء المتوسط لعينة التقنين؟ هل تقع فى مستوى أقل قليلا، أو أعلى قليلا من المتوسط؟ هل تقع فى الطرف الأعلى من التوزيع؟ وهكذا.

وحتى يمكن أن نحدد على وجه الدقة موضع الفرد بالنسبة لعينة التقنين فإن الاختبار النفسى يتحول إلى نوع من مقاييس الرتبة أومقاييس المسافة وتتحول الدرجة الخام إلى هذا النوع من القياس عن طريق البحث عن مقابلها من رتبة أو مسافة وتسمى هذه القيم المقابلة للدرجات الخام بالدرجات المشتقة derived scores ، وهذه الدرجات المشتقة تفيد فى تحقيق غرضين : أولهما تحديد الوضع النسبى للفرد فى العينة المعيارية وتقويم أدائه فى ضوء أداء الآخرين المماثلين له (أو المختلفين عنه)، وثانيهما أنها تعطينا مقاييس قابلة للمقارنة سواء بين الأفراد أو بين الاختبارات.

وحتى عهد قريب كانت المعايير هى الأساس الوحيد الذى يتطلبه علماء القياس النفسى للمقارنة بين الدرجات الخام وتفسيرها، حتى أنها كانت تعد شرطا جوهريا لجودة الاختبارات النفسية على اختلاف أنواعها وأغراض استخدامها. الا أنه مع التطور المعاصر فى النظرية الأساسية للاختبارات النفسية وظهور أنواع جديدة منها وأغراض متنوعة لها، وخاصة ما يسمى الاختبارات المرجعة إلى المحك توجه الباحثون إلى اعتبار المعايير أساسا واحدا من بين عدة أسس لهذه المقارنة، وأقتصر دورها على نوع واحد فقط من هذه الاختبارات هو ما يسمى «الاختبارات المرجعة إلى المعيار»، ونعرض فيما يلى المعايير الشائعة الاستخدام فى هذا النوع من الاختبارات.

المعايير الارتقائية أو النمائية :

تشير المعايير الارتقائية والنمائية developmental إلى المعنى الذى يعطى للدرجة الخام ويدل على اقتراب أو ابتعاد نمو المفحوص عن المسار الطبيعى لنمو العملية النفسية موضع القياس. وعلى الرغم من شيوع هذا النوع من المعايير (وخاصة فى صورة معيار العمر العقلى بالنسبة لاختبارات الذكاء) إلا أنه ليس دقيقا ولايصلح للمعالجة الإحصائية المحكمة ويجب ألا تتجاوز استخدامه حدود الأغراض الوصفية فى مجال البحث أو المواقف الكلينيكية ويوجد عدد من المعايير التى يستخدمها الأخصائى النفسى ومن هذه المعايير:

١ - **معايير العمر والفرق الدراسية :** لقد كان العالم الفرنسى ألفريد بينيه

Binet أول من استخدم مفهوم العمر العقلى Mental age فى تعديله لمقياسه المشهور فى الذكاء عام ١٩٠٨، وفى مقاييس العمر كما فى عينة التقنين، وتتطابق الدرجة العمرية التى يحصل عليها الفرد من أعلى مستوى عمرى يمكن أن يصل إليه بنجاح، فإذا استطاع طفل عمره الزمنى ١٠ سنوات أن يجيب بنجاح على أسئلة ١٢ سنة فإن عمره العقلى يصبح ١٢ سنة، ومعنى ذلك أنه يسبق عمره الزمنى بعامين، لأن أداءه يساوى أداء الطفل المتوسط من عمر زمنى أكبر منه بعامين.

وقد لوحظ أن أداء المفحوص فى الاختبارات التى تعتمد على معيار العمر العقلى يظهر قدرا من التشتت، وبعبارة أخرى فإن الطفل قد يفشل فى أداء بعض الأسئلة التى تقع فى مستوى عمرى أدنى من عمره العقلى، كما يمكنه الإجابة على بعض الأسئلة التى تقع فى مستوى أعلى منه،ولهذا السبب نلجأ عادة إلى حساب العمر القاعدى basal age ويقصد به أدنى مستوى عمرى يستطيع الطفل أن يجيب على جميع أسئلته والأسئلة السابقة عليه، ثم تضاف إليه درجات جزئية - فى صورة كسور من السنة أو شهور - تدل على جميع الأسئلة التى أجاب عليها من مستوى أعلى من العمر القاعدى وبذلك يصبح العمر العقلى للطفل هو مجموع العمر القاعدى والأشهر الإضافية التى تدل على ما حصل عليه الطفل من مستويات عمرية أعلى منه حتى يصل إلى سقف الاختبار والذى يدل على المستوى العمرى الذى يعجز الطفل عن الإجابة على جميع أسئلته والأسئلة اللاحقة عليه.

وقد لجأ إلى استخدام معايير العمر العقلى فى الاختبارات التى لاتنقسم مفرداتها إلى مستويات عمرية كما هو الحال فى اختبار ستانفورد - بينيه .، وفى هذه الحالة تكون الدرجة الخام هى عدد الإجابات الصحيحة أو الخاطئة على جميع أسئلة الاختبار - أو الزمن الكلى الذى استغرقه المفحوص فى حله وتحسب متوسطات الدرجات الخام التى تحصل عليها المجموعات العمرية المختلفة فى عينة التقنين، وتكون تلك المتوسطات هى معايير الأعمار العقلية للاختبار كما تستخدم معايير عمرية أخرى مثل العمر الطولى والعمر الوزنى وعمر التسنين والعمر الرسغى وعمر قبضة اليد وعمر القراءة وغير ذلك من الخصائص النفسية والجسمية ذات الصبغة الإرتقائية ويمكن بالطبع استخدام هذا الإجراء للحصول على **معايير الفرق الدراسية** فى حالة الاختبارات التحصيلية المقننة حيث تستخدم فئات الفرق الدراسية بدلا من فئات الأعمار.

ومن أهم مشكلات وحدة العمر العقلى ووحدة الفرقة الدراسية كمعيار اختبارى أنهما لا تتقدمان بمعدل ثابت فى مختلف مراحل العمر أو مراحل الدراسة، فالنمو العقلى أسرع وبالتالى يكون مقداره أكبر فى الطفولة المبكرة منه فى الطفولة المتأخرة، ويزداد بطئا وبالتالى يقل مقداره باقتراب الشخص من حد النضج، كما أن النمو التحصيلى ليس متكافئا فى جميع الفرق الدراسية ومعنى ذلك أن مقدار النمو الذى تتضمنه هذه الوحدات - بلغة القياس النفسى - يجعلها غير متساوية من عمر لآخر، وبالتالى يجعلها غير قابلة للمقارنة.

وقد تتضح هذه الفكرة إذا أعطينا مثالا من العمر الطولى، اننا فى هذه الحالة نجد أن الفرق بالسنتيمترات بين العمر الطولى للطفل فى سن ٣ أو ٤ أكبر من الفرق بين العمر الطولى فى سن ١٠، ١١، ومعنى ذلك أن عاما واحدا من «التفوق» أو «التخلف» فى سن ٤ سنوات يمثل اختلافا عن المستوى المتوسط أكبر كثيرا من مقدار يساويه من التفوق أو التخلف فى سن ١١ سنة.

٢ - **نسبة الذكاء التقليدية**: ابتكر علماء النفس معيار نسبة الذكاء in-telligence quotient أو I. Q للحصول على تفسير موحد للذكاء بصرف النظر عن العمر الزمنى للمفحوص، وقد أشار إلى أهمية ما يسمى النسبة العقلية شترن وتولمان، إلا أن هذا المعيار استخدم لأول مرة فى مقياس ستانفورد - بينيه الذى أعده ترمان عام ١٩٦١ ويمكن الحصول على نسبة الذكاء كما يلى:

$$\text{نسبة الذكاء} = \frac{\text{العمر العقلى}}{\text{العمر الزمنى}} \times ١٠٠$$

فإن تعادل العمر العقلى والعمر الزمنى تكون نسبة الذكاء فى هذه الحالة ١٠٠ وتدل على أداء متوسط أو عادى، وإذا قلت عن ١٠٠ دلت على تخلف أو ضعف، وإذا زادت عن ١٠٠ دلت على التفوق.

وتفيد نسبة الذكاء فى ضبط الضمور العقلى لطفل عمره الزمنى ٤ سنوات هو ٣ تصبح نسبة ذكائه فى هذه الحالة ٧٥ فإذا بلغ هذا الطفل سن ١٢ وكان عمره العقلى ٩ تظل نسبه ذكائه ثابتة أى ٧٥. أى أن نسبة الذكاء تدل على نفس الوضع النسبى للفرد فى الجماعة سواء فى سن ٤ سنوات أو فى سن ١٢ سنة، وبعبارة أخرى فإن نسبة

الذكاء يمكن استعمالها فى المقارنة فى مختلف الأعمار، بمعنى أن تفسير نسبة الذكاء يظل هو نفسه بصرف النظر عن العمر الزمنى للمفحوص.

ولابد أن نشير فى هذا الصدد إلى أنه لكى تظل نسبة الذكاء ثابتة فإن ذلك لايمكن - نظريا - إلا فى حالة واحدة فقط وهى أن وحدة العمر العقلى تتناقص (أى تصبح أصغر) تناقصاً يتناسب تناسبا مباشرا مع العمر الزمنى، وهذا شرط ضرورى حتى يمكن استخدام نسبة الذكاء استخداما له معنى.

وحين تتناقص وحدة العمر العقلى تبعا للتقدم فى العمر على هذا النحو فإن الفروق الفردية التى تقاس بهذه الوحدة يفترض فيها بالطبع أنها تتزايد بنفس النسبة، ويشبه ذلك من الناحية الحسابية أن الفروق الفردية فى الطول تكون أكبر ١٢ مرة حينما تقاس بالبوصات عنها لو قيست بالأقدام وذلك لأن البوصة هى $\frac{١}{١٢}$ من القدم، وبنفس المنطق فإن الفروق الفردية أوالتباين فى العمر العقلى يجب أن تكون فى سن ١٤ ضعفها فى سن ٧ سنوات ولايمكن لنسب الذكاء أن تظل ثابتة إلا إذا تزايد الانحراف المعيارى للأعمار العقلية بنفس نسبة زيادة الأعمار الزمنية، وبهذا الشرط وحده يمكن أن يكون لنسبة الذكاء نفس المعنى فى مختلف الأعمار.

وقد أمكن تحقيق هذا الشرط فى الطبعات الحديثة من اختبار ستانفورد - بينيه فقد اتضح من دراسة انتشار الأعمار العقلية من سن ٦ حتى سن ١٨ أن هناك اتجاها نحو زيادة التباين فى الأعمار العقلية تبعا لزيادة الأعمار الزمنية. ولكننا نجد فى كثير من الأحيان عدم تساوى الانحرافات المعيارية لنسب الذكاء فى مختلف الأعمار، وهذه الصعوبة أمكن التغلب عليها فى اختبار ستانفورد - بينيه بإعداد جدول تصحيح يستخدم فى تصحيح نسب الذكاء فى مستويات عمرية معينة.

وهناك مشكلة أخرى فى تطبيق معيار نسبة الذكاء نجدها فى اختبار البالغين والكبار فالواقع أن مفهوم معايير العمر على إطلاقه يدل دلالة مباشرة على أن فائدته مقصورة على الأطفال لما يتضمنه من معانى ارتقائية، أما بالنسبة للراشدين والكبار فإن ظواهر النمو ليست على نفس الدرجة من الوضوح إلا فى حالات تدهور الشيخوخة ولذا نجد أن فى اختبار مثل ستانفورد - بينيه لايتحسن أداء البالغ المتوسط كثيرا، صحيح أنه قد أضيف إلى الاختبار مستويات رفيعة لذكاء الكبار حتى يتوافر فيه

سقف مناسب وبالتالى يمكننا الحصول على أعمار عقلية أعلى، ولكننا إذا وجدنا شخصاً معينا يحصل على عمر عقلى مقداره ٢٠ فى هذا الاختبار فإن ذلك لايسمح لنا بتفسير مؤكد وعلى نفس الدرجة من الوضوح كما هوالحال فى الأعمار العقلية الدنيا، فمن المؤكد أن العمر العقلى ٢٠ لايمكن تحديده بما يمكن أن يؤديه الشخص المتوسط البالغ من العمر ٢٠ سنة، فى هذا الاختبار، أما إذا اختبرنا شخصاً بالغا ضعيف العقل يقل عمره العقلى عن حدود النضج فإن استخدام مفهوم العمر العقلى فى هذه الحالة عنه يتشابه باستخدامه مع الأطفال، أما فى حالة الراشدين والكبار من الأسوياء أو المتفوقين فلا مناص من استخدام أنواع أخرى من المعايير ليس من بينها معيار العمر لصعوبة تحديدهم عندهم.

أما المشكلة الثالثة التى تواجهنا حينما نفسر الدرجات الخام بدرجات العمر تتمثل فى أن هذا النوع من المعايير لايستخدم إلا فى الوظائف النفسية والعمليات العقلية التى تتغير تغييرا واضحاً ومتسقا مع النمو، أما بالنسبة للسمات التى لها علاقة ضئيلة بالعمر الزمنى أو لا تكون لها علاقة مباشرة به لايجوز بالطبع استخدام معيار العمر معها، ومن أمثلة ذلك مقاييس سمات الشخصية، فبعض هذه السمات يظل ثابتا فى معظم مراحل نمو الفرد.

ولأسباب تعود إلى هذه المشكلات وغيرها توقف علماء القياس النفسى المعاصرون عن استخدام نسبة الذكاء على النحو المشار إليه، وأصبح يطلق عليها اسم نسبة الذكاء التقليدية تمييزا لها من أنواع أخرى من المعايير أكبر دقة سنشير إليها فيما بعد ولعل فى هذا ما ينبه الباحثين فى مصر والبلاد العربية إلى أن يتوجهوا فى بناء الاختبارات النفسية وفى استخداماتها فى الأغراض المختلفة (وخاصة الأغراض الكلينيكية) إلى المعايير الأخرى المناسبة غير نسبة الذكاء التقليدية.

٣ – **المئينيات:** يعبر عن الدرجات المئينية percentiles فى صورة نسب مئوية للأفراد من عينة التقنين الذين يقعون أدنى من درجة خام معينة، وتدل على الوضع النسبى للفرد بمقارنته بعينة التقنين هذه، وهى تعد من نوع الترتيب فيما عدا أننا فى الترتيب نبدأ عادة باعتبار الشخص الذى يقع فى القمة ـ أى أفضل أفراد المجموعة من حيث الأداء ـ فى الرتبة الأولى، بينما فى المئينيات فإننا نبدأ العد من أسفل، وبالتالى فإن أدنى المئينيات يدل على أسوأ مكانة للفرد.

ويقابل المئينى الخمسون منتصف جماعة التقنين أو وسيطها فإذا زاد المئينى عن ٥٠ يدل هذا على أداء أعلى من المتوسط، وإذا قل عنه دل على أداء أقل من المتوسط، ويعرف المئينى ٢٥ ، والمئينى ٧٥ إحصائيا بالربيع الأدنى والربيع الأعلى من التوزيع، وهما مثل الوسيط يعطيانا معالم تصف توزيع الدرجات ومقارنته بغيره من التوزيعات، وهى كالوسيط والربيع فى طريقة حسابها، كما أنها تقسم التوزيع التكرارى إلى نقاط، وهى فى حالة المئينيات تصل إلى ٩٩ نقطة.

ويجب ألا نخلط بين المئينيات ودرجات النسب المئوية الشائعة، فدرجات النسب المئوية هى نوع من الدرجات الخام يعبر عنها فى ضوء نسب مئوية للإجابات الصحيحة. أما المئينيات فهى درجات مشتقة يعبر عنها فى ضوء نسب مئوية للمفحوصين فإذا افترضنا أن لدينا ٤٠ شخصا منهم ٢٧ حصلوا على درجات أعلى من الشخص (أ)، ١٢ حصلوا على درجات أقل منه، فإننا فى هذه الحالة نضع هذا الشخص بين المجموعتين معتبرين أنه تفوق على ٣١٪ من العدد الكلى للمجموعة، وتكون درجته المئينية، والدرجة الخام التى تقل عن أى درجة حصل عليها أفراد عينة التقنين تكون رتبتها المئينية صفرا، والدرجة الخام التى تزيد على أعلى درجة حصل عليها أفراد عينة التقنين تكون رتبتها المئينية ١٠٠ وهذه المئينيات لا تتضمن بحال من الأحوال أن تكون الدرجات الخام المقابلة لها مقدارها صفرا أو تدل على الأداء الكامل أوالحد الأقصى للدرجة الخام.

وللمئينيات ميزات عديدة، فهى سهلة الحساب ميسورة الفهم، كما أن استخدامها يصلح لكل أنواع الاختبارات، كما تصلح للأطفال والمراهقين والراشدين على حد سواء، إلا أن أهم مشكلاتها تنشأ من عدم تساوى وحداتها وخاصة عند أطراف التوزيع، فإن كان توزيع الدرجات الخام يقترب من التوزيع الاعتدالى فإن الفروق بين الدرجات الخام الأقرب إلى الوسيط أو مركز التوزيع تتزايد (أى تنتشر الحالات الكثيرة التى تقع فى المنتصف وتتفتت إلى مئينيات كثيرة) بينما تتضاءل هذه الفروق عند الأطراف (أى تتجمع الحالات القليلة التى تقع عند الأطراف بعضها إلى بعض فى عدد قليل من المئينيات)، فإذا أعدنا رسم التوزيع التكرارى بيانيا باستخدام المئينيات نحصل على ما يسمى التوزيع المستطيل.

ولابد أن نشير إلى أننا لا نستطيع استخراج متوسطات الدرجات المئينية، لأن هذا المتوسط يختلف فى مقداره عما لو استخرجنا المتوسط الحسابى للدرجات الخام ثم حولناه إلى درجة مئينية، وبدلا من اللجوء إلى المتوسط الحسابى لقياس النزعة المركزية للدرجات المئينية يمكن للباحث أن يعتمد على الوسيط باعتباره المقياس الإحصائى المناسب للنزعة المركزية فى هذه الحالة.

كما يجب الإشارة إلى أن الدرجات المئينية لايمكن المقارنة بينها ما لم تكن هذه المقارنة ممكنة من الجماعات التى تشتق منها فإذا اختلفت الجماعات، كأن تكون إحداها مثلا من تلاميذ مدرسة ثانوية صناعة والثانية من عمال تلمذة صناعية والثالثة من طلاب كلية للهندسة فإن الدرجات الخام التى تقابل مئينى معين فى كل حالة من هذه الحالات تختلف، كما أن الدرجة الخام الواحدة فى كل مجموعة من هذه المجموعات الثلاث تقابل مئينيات مختلفة ومن ذلك مثلا أن الدرجة الخام التى تقابل الدرجة المئينية الـ ٦٠ مثلا لطلاب الجامعة قد تقابل مثلا المئينى الـ ٧٥ فى مجموعة طلاب المدرسة الثانوية الصناعية وتقابل المئينى الـ ٩٠ فى مجموعة عمال التلمذة الصناعية، بعبارة أخرى إذا لجأنا إلى استخدام الدرجات المئينية لابد للباحث أن يضع فى اعتباره عينة التقنين التى تشتق منها هذه المعايير، ولو أن ذلك لايحول دون استخدام مجموعة من جداول التحويل إلى درجات مئينية تسمح للمختبر أو الباحث أن يقارن الشخص الذى يختبره بجموعات مرجعية مختلفة.

٤ - **الدرجات المعيارية**: يتزايد استخدام الدرجات المعيارية Standard scores كمعايير فى الاختبارات الحديثة، وتعتبر هذه المعايير أفضل صورة لتحويل الدرجات الخام لأنها تعبر عن بعد الدرجة الخام التى يحصل عليها المفحوص عن الدرجة الخام المتوسط منسوبا إلى الانحراف المعيارى للتوزيع، أى أن :

$$\text{الدرجة المعيارية} = \frac{\text{الدرجة الخام - المتوسط}}{\text{الانحراف المعيارى}}$$

$$\text{أو } ذ = \frac{س - م}{ع}$$

ويطلق على هذا النوع «الدرجات المعيارية الخطية Linear standard scores

ويقصد بذلك أننا نحتفظ بالعلاقات العددية الدقيقة للدرجات الخام الأصلية، فنحن نطرح من الدرجة مقدارا ثابتا (المتوسط)، ونقسمها على مقدار ثابت (الانحراف المعيارى)، ومعنى ذلك أن السعة النسبية للفروق بين الدرجات المعيارية التى تشتق من هذا التحويل الخطى تتطابق تماما مع الفروق بين الدرجات الخام، كما تتكرر جميع خصائص التوزيع الأصلى للدرجات الخام فى توزيع هذه الدرجات المعيارية.

ولعلنا نذكر أن من أهم أسباب تحويل الدرجات الخام إلى درجات مشتقة من المعايير عامة هو جعل الدرجات التى نحصل عليها من مختلف الاختبارات قابلة للمقارنة، وهذا الفرض لايمكن أن يتحقق للدرجات المعيارية الخطية إلا إذا كانت التوزيعات التكرارية للاختبارات المختلفة متشابهة، فإذا كان أحد التوزيعيين اعتداليا والاخر ملتويا فلابد فى هذه الحالة من إجراء تحويلات غير خطية للدرجات الخام حتى تتواءم مع نمط معين من أنماط التوزيع، وعادة ما يختار الباحثون فى مجال الإختبارات المرجعة إلى المعيار التوزيع الاعتدالى لأسباب نظرية خالصة، ومعنى ذلك أن التوزيعات المختلفة يجب أن تتحول فى هذه الحالة إلى توزيعات تتفق مع خصائص المنحنى الاعتدالى، ثم تحول الدرجات إلى درجات معيارية اعتدالية باستخدام جداول مساحات وارتفاعات المنحنى الاعتدالى normalized standard scores وتصبح لهذه الدرجات المعيارية الاعتدالية فى هذه الحالة نفس خصائص الدرجات المعيارية الخطية (أى متوسطها صفر وانحرافها المعيارى واحد صحيح).

إلا أن ما يجب أن ننبه إليه هو الدرجات المعيارية الخطية لا تتطابق مع الدرجات المعيارية الاعتدالية إلا فى حالة واحدة فقط وهى أن تكون التوزيعات التكرارية التى حسبت منها الدرجات المعيارية الخطية اعتدالية بالطبع، كما ننبه أيضاً إلى أن التحويلات الاعتدالية لا يجب أن تتم إلا إذا توافرت شروط الاعتدالية من حيث السمة المقيسة والعينة المختبرة والمقياس المستخدم، وهى شروط فصلها آمال صادق وفؤاد أبو حطب فى موضع آخر (١).

ومن الواضح أن حساب الدرجات المعيارية الخطية يتضمن أن تكون هذه الدرجات سالبة أو موجبة - فالدرجات السالبة تدل على أداء أدنى من المتوسط، والدرجات الموجبة تدل على أداء أعلى منه، والدرجة الخام التى تساوى المتوسط تقابل الدرجة المعيارية صفرا، كما أننا نلاحظ على هذه الدرجات المعيارية أنها

تتضمن كسورا عشرية حتى تتوافر لدينا معلومات كافية نميز بها بين مختلف المفحوصين وخاصة إذا علمنا أن المدى الكلى لأغلب المجموعات لا يتعدى نطاق ٣ انحرافات معيارية.

وللتغلب على هاتين الصعوبتين اقترحت عدة عديلات تهدف تحويل الدرجات المعيارية، إلى صورة أكثر ملاءمة وأيسر فى الفهم، وتعتمد هذه التعديلات على بعض الخصائص الرياضية للدرجات المعيارية، والتى تتمثل فى أن متوسط هذه الدرجات يساوى صفرا، وانحرافها المعيارى يساوى الواحد الصحيح، ولذلك نستطيع أن ندخل عليها ما نشاء من تعديلات باستخدام متوسطات وانحرافات معيارية جديدة، وفى تحويل الدرجات المعيارية الأصلية إلى المقياس الجديد يتطلب الأمر ببساطة أن نضربها فى الانحراف المعيارى الجديد ونجمعها إلى المتوسط الجديد، أى أن الدرجة المعيارية المعدلة تحسب كما يلى:

الدرجة المعيارية المعدلة = الدرجة المعيارية الأصلية × الانحراف المعيارى الجديد + المتوسط الجديد.

$$ذَ = (ذ \times ع_{َ}) + م_{َ}$$

وقد كثرت الدرجات المعيارية المعدلة وتعددت ، ويلخص الجدول رقم (١٥) بعض أنواعها الشائعة، وقد يكون أكثرها شيوعا فى الوقت الحاضر، وبخاصة فى اختبارات الذكاء، نسبة الذكاء الانحرافية، وهى تختلف تماما عن نسبة الذكاء التقليدية التى عرضناها آنفا، فنسبة الذكاء الانحرافية درجة معيارية معدلة متوسطها الجديد ١٠٠ انحرافها المعيارى الجديد قد يكون ١٥ (كما هو الحال فى الطبعات الحديثة من مقياس ستانفورد ـ بينيه) أو ١٦ (كما هو الحال فى اختبارات وكسلر).

وقد تأثر بناة اختبارات الذكاء بمفهوم نسبة الذكاء فى معيارهم القائم على فكرة الدرجة المعيارية (نسبة الذكاء الانحرافية) لأسباب تاريخية ترجع فى جوهرها لاستخدام المصطلح لفترة طويلة فى علم النفس، واختير الانحراف المعيارى الجديد (١٦ أو ١٥) فى ضوء التحليل الإحصائى للانحرافات المعيارية لنسب الذكاء التقليدية والتى بلغ وسيطها فى مقياس ستانفورد ـ بينيه طبعة ١٩٣٧ حوالى ١٦، أما المتوسط ١٠٠ فقد اختير لأنه يشير فى نسبة الذكاء التقليدية إلى المتوسط، وهكذا تتوافر فى

نسبة الذكاء الانحرافية (كدرجة معيارية معدلة) جميع الخصائص «الشكلية» لنسبة الذكاء التقليدية إلا أنها تختلف عنها تماما في الأسس الرياضية وطرق الحساب كما بينا بالإضافة إلى اختلافهما فى المعنى التفسيرى فنسبة الذكاء التقليدية هى نسبة العمر العقلى إلى العمر الزمنى، أما نسبة الذكاء الانحرافية فتدل على اقتراب أو ابتعاد عن المتوسط بمسافات من الانحراف المعيارى.

جدول (١٥) بعض أنواع الدرجات المعيارية المعدلة

المتوسط	الانحراف المعيارى	الدرجة المعيارية التى تقابل انحرافا معياريا واحدا أعلى من المتوسط	الدرجة المعيارية التى تقابل انحرافين معياريين أقل من المتوسط	اسم الدرجة المعيارية
صفر	١	١	-٢	الدرجة المعيارية الأصلية
٥	٢	٧	١	الدرجة التساعية أو التساعى المعيارى
٥٠	١٠	٦٠	٣٠	الدرجة التائية
١٠٠	١٥ أو ١٦	١١٥ أو ١١٦	٧٠ أو ٦٨	نسبة الذكاء الانحرافية

تعليق على المعايير: المعايير إذن هى أسس للحكم على أداء المفحوصين والمقارنة بينهم فى ضوء أدائهم الفعلى فى الاختبارات ولأنها ذات طبيعة كمية يمكن استخدام الحاسبات الالكترونية معها بحيث يقدم لنا الكمبيوتر تفسيرا للدرجات الخام، وحتى يمكن المقارنة بين المعايير الخاصة بالاختبارات المختلفة، أو بين المعيار الواحد مع المفحوص فى مراحل مختلفة من تطبيق الاختبار ، ولابد من أن يكون الاختبار المستخدم والذى تشتق من درجات عينة تقنين المعايير الخاصة به معلوما للفاحص. فالمعايير لا تستخدم لأغراض المقارنة المجردة المطلقة.

وتفيدنا معرفة طبيعة الاختبار المستخدم فى القياس فى تحديد محتواه وما يقيسه، ونود أن ننبه هنا إلى الاختبارات التى تبدو لنا متشابهة فيما تقيس (كإختبارات الذكاء) هى فى الواقع مقاييس لعمليات نفسية مختلفة.

وحتى يدرك الفاحص مغزى المعايير ومعناها عليه أيضًا أن يتحقق من القيم

العددية المستخدمة فى حساب المعايير فحتى يستطيع المقارنة بين نسب الذكاء الانحرافية لاختبارين للذكاء لابد أن تكون واضحة تماما وحدات القياس المستخدمة فى حساب هذا المعيار ـ فنسب الذكاء الانحرافية (المتوسط جديد مقداره ١٠٠) تختلف تبعا للقيم المختارة للانحراف المعيارى الجديد (١٥ أو ١٦ مثلا).

ولكى يفهم الفاحص طبيعة المعيار ويستخدمه استخداما صحيحا لابد أن يتعرف بعناية على عينة التقنين المستخدمة فى حساب معايير الاختبار، وعلينا دائما أن نعلم أن معايير أى اختبار لابد أن تتجاوز حدود خصائص عينة تقنينه بشرط أن تتوافر فيها دائما شروط العينة الجديدة وخاصة شرط التمثيل ، فإذا أردنا استخدامه مع مفحوصين ذوى خصائص مختلفة لابد من إعادة تقنينه على عينات جديدة، وهنا تجدنا أزاء موقف هام بالنسبة للمعايير، وهو أن الاختبار الواحد قد تكون له معايير متعددة بعضها قد يكون عاما على المستوى القومى أو أقل عمومية على المستوى المحلى، بل قد يصل الأمر إلى حساب معايير خاصة لعينات تقنين مختلفة تشتق من مختلف الفئات التى يصنف إليها المجتمع الواحد أو الثقافة الواحدة، وفى هذه الحالة ينسب المفحوص إلى الأصل الإحصائى الذى ينتسب إليه ويكون الحكم على أدائه وتفسيره أقرب إلى العدالة.

صحيح أنه ظهرت فى السنوات الأخيرة بعض الاتجاهات العملية والنظرية تعين على بناء مقاييس موحدة «متحررة من أثر العينة» وتصلح للتطبيق على الأفراد والجماعات من مستويات مختلفة فى القدرة ولقياس محتوى من مستويات مختلفة من الصعوبة، ومن هذه الطرق ما يسمى استخدام المجموعة المرجعية الثابتة fixed ref-erence group ونماذج السمة الكامنة latent trait models ، وهى تعتمد على مجموعة من الأسئلة أوالاختبارات الفرعية تسمى أسئلة أو اختبارات الارتكاز anchor وتختاز بمواصفات معينة بحيث تتجاوز العينات والاختبارات والأسئلة، إلا أن هذه الطرق جميعا تهتم بأحد مصادر الفروق بين الأفراد والجماعات هو صعوبة الأسئلة ، وتوجد مصادر أخرى لا تقل أهمية.

وتبقى أخيراً الإشارة إلى أن المعايير كأسس للحكم على الدرجات الخام وتفسيرها تصلح للاختبارات التى تسمى فى الوقت الحاضر الاختبارات المرجعة إلى المعيار دون سواها، ولا تصلح للأنواع الأخرى التى تستخدم معها المحكات والمستويات كما سنبين فيما بعد.

المحكات:

يمكن للدرجات الخام أن تفسر أيضاً فى ضوء الأداء فى محك متوقع، والمحك كما أشرنا من قبل هو أساس خارجى مستقل للحكم على الأداء فى الاختبار، كما هو الحال فى برامج التدريب أو مقاييس الكفاية التعليمية والمهنية، ويتفق هذا المعنى للمحك مع استخدامه فى بحوث الصدق، وإلى هذا المعنى يجب أن يشير ما يسمى فى الوقت الحاضر الاختبارات المرجعة إلى المحك إلا أن المعنى الاصطلاحى الشائع لهذا النوع من الاختبارات يخرجها من المعنى الحالى للمحك ويدخلها فى فئة جديدة سنشير إليها تحت اسم «المستوى» وعلى ذلك فإن أسس الحكم على ما يسمى «الاختبارات المرجعة إلى المحك» ليست المحكات كما شاعت فى مجال القياس النفسى، وإنما هى المستويات، وكان الأجدى لها أن تسمى «الاختبارات المرجعة إلى المستويات» حتى لا توقعنا فى هذا الخلط السيمانتى الذى لا مبرر له . وعلى كل فقد أثرنا الإشارة إليها بالاصطلاح الذى شاع فى الوقت الحاضر على أن يكون القارئ واعيا طول الوقت فى أن ما يقصد بالمحك فى هذا النوع من الاختبارات هو المستوى.

أما المحكات حين تستخدم فى الحكم على الأداء الاختبارى وتفسيره، فإن ذلك توسيعا لمعناها فى بحوث الصدق وامتدادا له إلى عالم الواقع، ولعل ما يتفق مع هذا المعنى ما اقترحه فؤاد أبو حطب من إشارة إلى الصدق المرتبط بالمحكات بالصدق العملى التطبيقى(*)، ويتضح ذلك من عرض أشهر الطرق التى تستخدم فيها المحكات فى الحكم على الدرجة الخام فى الاختبارات النفسية وتفسيرها وهى ما يسمى طريقة جداول التوقع Expectancy tablas.

جداول التوقع: سبق أن قلنا أن كلا من الصدق التنبؤى والصدق التلازمى يعتمد على تقويم درجات الاختبار فى ضوء محك مستقل يهدف الاختبار إلى التنبؤ به أو تشخيصه على التوالى، والعلاقة بين درجات المفحوصين فى الاختبار ومكانتهم فى المحك يمكن أن يحللها الباحثون بطرق مختلفة، وأبسط صورة للتعبير عن هذه العلاقة هى ما يسمى جدول التوقع ومثل هذا الجدول يوضح توقع احتمال حصول الفرد على

(*) استخدم فؤاد أبو حطب هذا المصطلح فى سلسلة محاضراته التى ألقاها عامى ١٩٧٩ – ١٩٨٠، ١٩٨٠ – ١٩٨١ بالمركز القومى للبحوث الاجتماعية والجنائية بالقاهرة عن الجوانب السيكومترية للاختبارات الإسقاطية.

درجة معينة فى المحك (وليكن النجاح المدرسى أو المهنى) من حصوله على درجة معينة فى الاختبار.

وقد يأخذ جدول التوقع صورة رسم بيانى للمحك ذى البعدين (نجاح أو فشل مثلا) يوضح العلاقة بين النسب المئوية للأشخاص الذين استبعدوا من التدريب المبدئى على الطيران وبين الدرجة التساعية التى يحصل عليها الشخص فى بطارية انتقاء الطيارين التى أعدها السلاح الجوى الأمريكى، ويلاحظ من هذا الجدول أن ٧٧٪ من الأشخاص الذين حصلوا على التساعى الأول استبعدوا بينما نجد أن حوالى ٤٪ فقط من أولئك الذين حصلوا على التساعى التاسع فشلوا فى إتمام التدريب. كما يتضح لنا من الجدول أن النسب المئوية تتزايد تدريجيا فى التساعيات التى تقع بين الأول والأخير.

والنسب المئوية المعطاة فى مثل هذا الرسم تعطى تقديرا لاحتمال حصول الأفراد الذين تختبرهم على مكانة معينة فى المحك، فنحن نتنبأ مثلا من هذا الشكل أن حوالى ٤٠٪ من طلاب الطيران الذين يحصلون على الدرجة التساعية ٤ سوف يفشلون وأن حوالى ٦٠٪ منهم سوف ينهون برنامج التدريب على أعمال الطيران بنجاح ويمكن تقرير توقعات مشابهة بنجاح أو فشل الأفراد الذين يقعون فى أى من المعايير التساعية. وعلى ذلك فالفرد الذى يحصل على درجة تساعية مقدارها ٤ تكون فرصته $\frac{٦٠}{٤٠}$ أى $\frac{٣}{٢}$ فى انهاء برنامج التدريب على أعمال الطيران بنجاح.

وعندما يكون المحك متصلا وليس من النوع ذى البعدين فحسب فان جدول التوقع يمكن اعداده اعدادا مباشرا من الرسم البيانى للانتشار، وفى اعداد رسم الانتشار يحدد لكل فرد درجتان فى الاختبار وفى المحك وفيه يمثل الخط الأفقى مقاييس المحك (تقديرات التلاميذ فى مادة البلاغة) ويمثل الخط الرأسى الدرجات فى اختبار الجمل. وفى هذا الاختبار يطلب من المفحوص أن يحدد الأخطاء فى النحو أو علامات التوقيف أو استعمال الكلمات فى سلاسل من الجمل. وكل علامة فى الشكل تدل على كل من الدرجة فى الاختبار والتقدير فى المحك لكل طالب من مائة، كما أن الرسم يتضمن العدد الكلى لتكرارات فى الخلايا المختلفة وكذلك المجموع الخام للتكرارات.

ولتحويل جدول الانتشار هذا الى جدول توقع تحول التكرار الموجود فى كل خلية

الى نسبة مئوية من المجموع الخام المقابل له. فمثلا يوجد ١٦ شخصا حصلوا على الدرجات بين ٣٠ و٣٩ فى الاختبار، من بينهم ٦٪ (شخص واحد) حصل على التقدير (هـ) فى مقرر البلاغة و١٩٪ (٣ أشخاص) حصلوا على التقدير (د) و٥٦٪ (٩ حالات) حصلوا على التقدير (جـ) و١٩٪ (٣ أشخاص) حصلوا على التقدير (ب). وبنفس الطريقة يمكن أن تحول جميع تكرارات الخلايا فنحصل على جدول التوقع الآتى:

جدول (١٦)

جدول توقع يوضح العلاقة بين الدرجات فى اختبار الجمل وبين تقديرات مادة البلاغة

درجات الاختبار	النسب المئوية للافراد الذين حصلوا على كل تقدير				
	هـ	د	جـ	ب	أ
٨٠-٨٩					١٠٠
٧٠-٧٩				٢٠	٨٠
٦٠-٦٩			١٤	٦٣	٢٣
٥٠-٥٩			٣٩	٣٥	٢٦
٤٠-٤٩		١٤	٥٩	٢٧	
٣٠-٣٩	٦	١٩	٥٦	١٩	
٢٠-٢٩	١٣	٥٠	٣٧		
١٠-١٩		١٠٠			
صفر - ٩		١٠٠			

ومن هذا الجدول يمكن التنبئو بالتقدير الذى يحصل عليه الفرد فى المحك من الدرجة التى يحصل عليها فى الاختبار. فمثلا يمكن أن نتوقع أن الشخص الذى يحصل على درجة أعلى من ٨٠ أن يحصل على تقدير (أ) والذى يحصل على درجة أعلى من ٧٠ أما أن يحصل على تقدير (أ) أو (ب)، كما نتوقع أن الشخص الذى يحصل على درجة أعلى من ٥٠ لن يحصل على تقدير (هـ) أو (د). وهناك عديد من التوقعات التى يمكن الوصول اليها من هذا الجدول.

٠لابد أن نشير فى هذا الصدد أن مثل هذه التنبؤات تتعرض لاخطاء العينة

وخاصة إذا كان عدد الأفراد الذين نختبرهم صغيرا. وحيث أن النسب المئوية الفردية تعتمد على عدد صغير نسبيا من الأفراد فى كل خلية فإن ذلك يعنى أن التغيرات الناتجة عن الصدفة عند الانتقال من عينة لأخرى قد تكون كبيرة.

المستويات:

النوع الثالث من الأسس التى تستخدم فى الحكم على الأداء الاختبارى وتفسيره ما يسميه آمال صادق وفؤاد أبو حطب (٩) المستويات standards وهى - كما ذكرنا آنفا تتشابه مع المعايير فى أنها أسس للحكم على الأداء فى ضوء هذا الأداء ذاته وليس فى ضوء أسس خارجية عنه، إلا أنها تختلف عن المعايير فى أنها قد تأخذ الصورة الكمية أوالكيفية، وفى أنها تتحدد فى ضوء ما يجب أن يكون عليه الأداء الاختبارى وليس ما هو عليه بالفعل (كما هو الحال فى المعايير).

ولعل هذا النوع هو ما يقصده بناة ما يسمى فى الوقت الحاضر الاختبارات المرجعة إلى المحك «فالمحك» المستخدم فى هذه الاختبارات ليس أساسا «خارجيا» وإنما هو أساس «داخلى» من بنية الاختبار ذاته ولعل أول مع أوقع حركة الاختبارات النفسية فى هذا الخلط روبرت جليزر R. Glaser ، أول من استخدم هذا المصطلح عام ١٩٦٣ دون أن يتنبه إلى استخدام مصطلح «محك» فى القياس النفسى بمعنى ثابت ومحدد. ومما يؤسف له حقا أن سار على منواله عدد كبير من قادة القياس النفسى أنفسهم، وحينما وجدوا أنفسهم فى حيرة من أمرهم حاولوا البحث عن حلول توفيقية لعل أشهرها محاولة أنستازى فى طبعة ١٩٨٢ من كتابها الشهير(١٣ : ٩٤ - ١٠١).

ونعرض فيما يلى استخدام المستويات فى أنواع مختلفة من الاختبارات النفسية:

١ - **مستوى نطاق المحتوى:** لعل أقوى ما يدعم وجهة نظرنا فى استخدام مصطلح مستوى بدلا من مصطلح محك كأساس تفسيرى للأداء فى «الاختبارات المرجعة إلى المحك» ما تتطلبه طبيعة هذه الاختبارات ومعظمها من نوع الاختبارات التحصيلية من الرجوع - فى الحكم والتفسير على الأداء الاختبارى. إلى نطاق المحتوى content domain بدلا من الرجوع إلى خصائص عينة تقنين معينة (كما هو الحال فى المعيار) أو إلى خصائص أداء مستقل عن الاختبار (كما هو الحال فى المحك)، ولهذا يشار إليها أحيانا بالاختبارات المرجعة إلى المحتوى أو النطاق أو الأهداف، وهذا ما

التقويم النفسى ـــــــ

بحوث علم نفس النمو والتى أدت إلى وصف السلوك الإنسانى فى مراحله التتابعية ومن أشهر الأعمال المبكرة فى هذا المجال جهود جيزل وزملائه فى فى جامعة ييل بالولايات المتحدة والتى نتجت عنها جداول جيزل الارتقائية Gesell Develop-mental Schedules (١٤: ٥٧) وتوضح هذه الجداول المستوى النمائى التقريبى التى يصل إليه الطفل مقدرا بالشهور فى أربعة مجالات رئيسية هى النشاط الحركى والتوافقى واللغوى والشخصى الاجتماعى. ويتحدد هذ المستوى بمقارنة سلوك الطفل وبالسلوك المعتاد للأطفال من سن ٤ أسابيع وحتى ٣٦ شهرا ونشأت عن ذلك مقاييس ذات طبيعة رتبية متضمنة ثلاثة افتراضات أساسية هى:

١ – يتسم السلوك المبكر فى نموه بالتتابع المنتظم.

٢ – تتتابع المراحل النمائية الارتقائية لسلوك الإنسان بنظام ثابت.

٣ – تعتمد كل مرحلة لاحقة على إتقان الأنماط السلوكية اللازمة للمراحل السابقة.

ومنذ مطلع الستينيات شهد الميدان اهتماماً كبيراً بالنظريات الارتقائية لعالم النفس السويسرى العظيم جان بياجيه، والتى توصل إليها من خلال برنامج طموح للبحث فى النشاط المعرفى عند الأطفال والمراهقين بدأه منذ منتصف العشريات، وقد وظفت المهام التى استخدمها فى معظم الدراسات اللاحقة فى علم نفس النمو، كما بنيت اختبارات عقلية مقننة على أساسها وتعد هذه الاختبارات من النوع الرتبى أيضاً وتصدق عليها الافتراضات السابقة.

وللتمييز بين هذا النوع من المعايير ومعيار العمر العقلى نشير إلى أنه فى المعايير الرتبية تصمم المهام (التى قد تتحول إلى أسئلة أومفردات فى الاختبار) بحيث تكشف عن الجوانب المهمة والسائدة لكل مرحلة ارتقائية، ثم تجمع بعد ذلك البيانات الامبريقية حول الأعمار التى يتم التوصل فيها إلى كل مرحلة من هذه المراحل، أما فى معيار العمر العقلى فالعكس هو الصحيح، ففيه تختار المهام (الأسئلة) على أساس تمييزها بين الأعمار المتتابعة، ويوجد وجه آخر للاختلاف بين نوعى المعايير هو أن معيار العمر العقلى كمى فى جوهره ويتحدد فى ضوء درجة يحصل عليها المفحوص فى الاختبار، أما فى المعيار الارتقائى الرتبى فإن الأهم هو الوصف الكيفى لسلوك الطفل ولذلك فهو يعتمد على المستوى وليس المعيار.

التقويم النفسى ـــــــ

13. Anastasi, A, Psychological Testing (3rd. ed) New York: Macmillan, 1968, 1976, 1982.
14. Anastasi. A, Urbina, S. Psychological Testing (7th edit.) Prentice Hall, 1997.
15. A. P. A. Ethical Standards for Psychologists, Washington, D. C. 1953.
16. A. P. A. Technical Recommendation for Psychological Tests and Diagnostic Techniques, Washington, D. C. 1954.
17. A. P. A. Standards for Educational and Psychological Tests and Manuals. Washingon, 19 , D. C. 1966. 1974.
18. A. P. A. Ethical Standards for Psychologists, Amer Psychologist,. 1968, 23, 357 - 361.
19. _____ Casebook on Ethical Statements for Psychogists Washington, D. C., 1974.
20. _____ Ethical Principles in the Conduct of Research with Human Participants, Washington, D. C. 1973.
21. Benton A. L. Influence on Incentives Upon Intelligence Test Scores of School Children, J. Genet. Psychol, 1936, 49, 494, 496.
22. Borden, E, S, Psychological Counseling, New York, Appleton, 1965.
22a. Cohen, R., Moutague, P., Nathanson I., & Swerdlike, M., 1988: Psychological Testing an introduction to Tests and Measurement. California, Mayfield Publishing co. p187-188.
23. Cronbach, J. L. Essentials of Psycholoical Testing (3rd ed.) New York: Harper, 1960, 1970, 1984.
24. Cronbach, J.L., Meehi, P. E. Construct Validity in Psychological tests. Psychol, Bull., 1955, 52, 281 - 302.
25. Flanagan, J. C., The development of an Index of Examine Motivation Educ, Psychol. Measurement, 1955, 15, 144 - 161.
26. Ghiselli, G. E., Campbell J. P, ans Zadeck, S. Measurement theory for the Behavioural Sciences, San Francisco: W. H. Freeman, 1981.

27. Guilford, J. P. Psychometric Methods, New York: McGraw-Hill, 1954.
28. Hurlock, E. B. An Evaluation of Certain Incentives Used in School Work, J. Educ. Psychol., 1925, 16, 154 - 159.
29. Lindquist, E. F. (ed.) Educational Measurement Washington, D. C., Amer. Council on Educ. 1951.
30. Nunnally, J. C. Psychometric Theory, New York: McGraw-Hill, 1978.
31. Ronan, W. W. & Prien, E. P. Perspectives on the Measurement of Human Performance, NewYork: Appleton-Century-Crofts, 1971.
32. Sarason, S. B. et al., A Test Anxiety Scale for Children, Chlid. Develp., 1959, 29, 105 - 115.
33. Stanley, J. C. & Hopkins, K. D. Educational and Psychological Measurement Evaluation, N. J.: Prentice-Hall, 1972.
34. Thorndike, R. L. (ed.) Educational Measuremnt (2nd ed.) Washington: D. C. Amer. Council on Educ., 1971.
35. Thorndike, R. L. & Hafen, E. Measurement and Evaluation in Psychology and Education (4th ed.) NewYork: Wiley, 1977.

الـباب الثالث
اختبارات الذكاء العام

يشمل هذا الباب الفصول التالية:

الفصل السابع: اختبارات الذكاء الفردية

الفصل الثامن : اختبارات الذكاء الجماعية

تقديم الباب الثالث

يتناول الباب الثالث بالعرض فى الفصل السابع اختبارات الذكاء العام الفردية، وفى الفصل الثامن الإختبارات الجماعية، ثم رؤية لمفهوم شامل للذكاء اقترحه فؤاد أبو حطب تم عرضه فى الفصل التاسع. واختبارات الذكاء التى تم عرضها تعكس الإتجاه السيكومترى أو النمائى أو نظريات تجهيز المعلومات، وقد يكون الإتجاه الأول هو الأكثر شيوعا فى بناء اختبارات الذكاء العام أو القدرة العقلية العامة.

الإتجاهات النظرية فى قياس الذكاء

ظهر العديد من الاتجاهات النظرية فى تحديد العوامل المتمثلة فى الذكاء العام، أو العامل العام والقدرات الخاصة والتحصيل ويمكن أن نرجع هذا الإتجاه إلى بينيه فى فرنسا، ترمان فى الولايات المتحدة. واعتمد تصور العلماء الممثلين لهذا الإتجاه عن طبيعة الذكاء مستخدمين فى ذلك منهج التحليل العاملى والذين توصلوا إلى نتائج تؤكد ظهور عوامل مسئولة عن الإختلافات بين الناس. وظهر فى التراث السيكولوجى المفاهيم الأساسية التى ارتبطت بهذه الإتجاه مثل العامل العام والعوامل الخاصة التى أكدها سبيرمان فى نظريته كما قدم مجموعة من القوانين التى تفسّر النشاط العقلى أطلق عليها القوانين الإبتكارية (قانون إدراك العلاقات والمتعلقات) أو العامل العام المسئول عن الاداء في مختلف المواقف التى تتطلب نشاطا عقليا، أما العوامل الخاصة فهى مسئولة عن الأنشطة العقلية الخاصة.

وقد أدى هذا النموذج إلى ظهور الاختبارات التى تقيس الذكاء العام مثل إختبار المصفوفات المتتابعة لرافن والذى سيلى عرضه تفصيلا.

وفى مقابل هذا الإتجاه قدم ثرستون نموذجه عن الذكاء بإعتبار أن العقل البشرى يعمل من خلال مجموعة من القدرات الأولية حيث توصل فى دراساته إلى عدد من العوامل المنفصلة مع عدم إفتراض العامل العام، وقد ساعد فى الوصول إلى نتائج مغايرة إختلاف طرق التحليل العاملى وكانت من أهم هذه الطرق نجاحا طريقة هوتلنج فى المكونات الأساسية (فؤاد أبو حطب ٩٦ : ٨٠) والتى شاع استخدامها مع انتشار الحاسبات الالكترونية فى الوقت الحاضر والطريقة المركزية التى ابتكرها ثرستون وكثر شيوعها فى الإستخدام لسهولتها النسبية، ومن خلال ذلك تحددت معالم

نموذج ثرستون (العوامل المتعددة) ، أما العامل العام فى هذا النموذج فإنه يمثل عاملا من الدرجة الثانية حيث أخضع ثرستون مصفوفة إرتباط العوامل الأولية للتحليل العاملى بالطريقة المركزية وتوصل إلى العامل العام من الدرجة الثانية حيث يدل على القدرة المشتركة بين جميع القدرات العقلية أو الذكاء العام . فؤاد ابو حطب (٩٦ :٨٤) .

وبناء على هذا النموذج ظهر اختبار القدرات العقلية الأولية وأجريت عليه العديد من الدراسات والتى سيتم عرضها تفصيلا .

وقد أثار نموذج العوامل المتعددة لثرستون إهتماما كبيرا بميدان القدرات العقلية الأولية للتحقق من العوامل الأولية التى تم التوصل إليها من ناحية وتحليلها إلى مكوناتها البسيطة من ناحية أخرى (فؤاد أبو حطب ٩٦ :٨٧) وقد أدى ذلك إلى زيادة مضطردة فى عدد القدرات العقلية حتى بلغ أكثر من مائة قدرة عقلية (Anatasi 1990) وأدى هذا إلى الحاجة إلى تنظيم الميدان والذى تطلب ظهور بعض النماذج النظرية الجديدة التى تسعى لتوضيح العلاقات بين مختلف القدرات ومن أهمها النموذج الهرمى التراتبى والذى يعتمد على التعرف على الفئات، وعلى الفئات داخل الفئات، ويترتب على ذلك أن يصبح أسلوب التصنيف كالشجرة المعكوسة جذرها إلى أعلى وأغصانها إلى أسفل وفيه إفتراض وجود مستويات عديدة من العوامل (فؤاد أبو حطب ٩٦ :٨٨)

وقد أنتقل النموذج الهرمى فى علم النفس من ميدان العلوم البيولوجية ويعود فى أصله إلى هربرت سبنسر. وقد شاعت هذه الفكرة عند عدد من علماء النفس البريطانيين، وقد كانت بداية النماذج الهرمية فى نموذجى سيريل بيرت C.Burtوفيليب فرنون P.E. Vernon . وقد أسهم علم النفس المصرى منذ وقت مبكر فى ظهور نموذج فؤاد البهى السيد، وهو نموذج هرمى متكامل للقدرات العقلية المعرفية (فؤاد البهى – الذكاء – ١٩٧١) (حيث ميز فيه القدرة المعرفية العامة – القدرات الطائفية الكبرى – القدرات الطائفية المركبة – القدرات الطائفية البسيطة– القدرات الخاصة) أما أكثر النماذج الهرمية شيوعا فى الوقت الحاضر هو نموذج ريموند كاتل.

وقد حاول كاتل التوفيق بين نموذجى سبيرمان وثرستون حيث نشر مقالا شهيرا

عام ١٩٤١ حيث إفترض وجود بنية هرمية للقدرات العقلية، فالعامل العام يوجد فى جميع مقاييس القدرة بينما القدرات الأولية تنتمى إلى مستوى أدنى فى التنظيم الهرمى، وأن قدرات المستوى الأعلى تتحدد بمجموعة أكثر إتساعا من التغيرات إذا ما قورنت بالقدرات الفردية التى تتضمنها.

وظت أفكار كاتل تنمو وتتغير وتتطور وتتعدل حتى نشر فى عام ١٩٦٣ دراسة حاسمة (Cattell 1963) توصل فيها إلى نتيجة هامة وهى قابلية العامل العام عند سبيرمان إلى القسمة إلى عاملين من الدرجة الثانية أطلق على إحداهما الذكاء السائل (Flued) وعلى الآخر الذكاء المتبلور (Crystallized) ويتحدد الذكاء السائل بالاختبارات التى يفترض أن تقيس الوسع البيولوجى Biological Capapility لدى الفرد على إكتساب المعرفة وأهم هذه الاختبارات تلك التى تقيس عاملى الإستدلال الإستقرائى والقدرة المكانية عند ثرستون. وقد طور كاتل إختبارات للذكاء متحررة من أثر الثقافة تعرض على المفحوص مشكلات إستدلالية جديدة تتطلب إستخدام عناصر الخبرة المشتركة وإفترض أنها مقاييس جيدة للذكاء السائل بإعتباره إنعكاسا لقانونى إدراك العلاقات وإدراك المتعلقات عند سبيرمان (فؤاد أبو حطب ٩٦: ١٠٣).

أما الذكاء المتبلور فيتحدد بالإختبارات التى يفترض فيها أنها تقيس آثار التدريس Schoolingوالتثقيف Aculturation وأهم هذه الاختبارات تلك التى تقيس العامل اللفظى عند ثرستون.

وتختلف وجهة نظر بياچيه عن الذكاء اختلافا جوهريا عن تلك التى حددها أصحاب النظريات العاملية، والتى توصل لها من خلال ملاحظاته لسلوك الأطفال واستجاباتهم فى مواقف التفكير والإستدلال. وتوصل من خلال هذه الملاحظات أن هناك نظام عقلى يحكم تفكير الأطفال والذى يختلف عن تفكير الكبار، ومن خلال ملاحظاته أيضا- والتى استمرت لأكثر من نصف قرن - أمكن تحديد التراكيب المعرفية التى تصف وتفسر العمليات العقلية.

وقد ظهر العديد من الإختبارات التى اعتمدت على نظرية بياچيه ومن أهمها المقاييس البريطانية للقدرات .B.A.S والذى تضمنت عددا من اختبارات بياچيه وبناء على هذه الإختلافات النظرية توجد فئة من الاختبارات النفسية تسمى اختبارات

الذكاء العام، وتتميز هذه الاختبارات بأنها تستخدم فى مواقف متنوعة ويتحدد صدقها فى ضوء محكات أكثر شمولا، كما تتميز هذه الاختبارات بأنها تعطى درجة واحدة مثل نسبة الذكاء لتدل على المستوى العقلى العام للمفحوص، ويرى ترمان وميرل (٥٦) أن الوصول إلى هذا التقدير الشمولى للمستوى العقلى فإنه يتطلب خبرة من الباحث فى «تصويب الرماح إلى الأهداف الحاسمة»، ولتحقيق ذلك عليه أن يعرض على المفحوص عدداً كبيراً ومتنوعا من الأعمال على أساس أنها تمثل عينة ملائمة من أساليب الأداء العقلى، ومع ذلك فإن البحوث الحديثة تؤكد أن هذه الاختبارات تتشبع بعمليات عقلية معينة مثل القدرة اللغوية، وتتجاهل تماما بعض العمليات العقلية الأخرى، كالقدرة الموسيقية على سبيل المثال.

وكثيرا ما تسمى اختبارات الذكاء العام باختبارات الاستعداد المدرسى Scholasic aptitude لأن صدقها يتحدد فى العادة فى ضوء محكات التحصيل الأكاديمى. كما أن هذه الاختبارات تستخدم فى أغلب الأحوال فى أغراض التصفية المبدئية العامة (سواء من الناحية التربوية أو المهنية أو العسكرية)، والمفروض أن يتبعها تطبيق اختبارات الاستعدادات الأكثر تخصصاً، ويشيع هذا على وجه الخصوص فى اختبار المراهقين والراشدين الأسوياء عند الإرشاد أو الانتقاء أو التصنيف أو التوجيه أو غير ذلك من الأغراض التى أشرنا إليها فى الفصل الأول من هذا الكتاب، كما يوجد استخدام آخر لاختبارات الذكاء، فى الأغراض الكلينيكية خاصة، وهو تحديد الضعف العقلى وتشخيصه، وفى مثل هذه الأحوال فإن استخدام الاختبارات الفردية أكثر شيوعا من غيره ، إلا أننا كثيرا ما نستخدم الاختبارات الجماعية للذكاء فى أغراض البحث والممارسة السيكولوچية.

وسوف نخصص فى هذا الباب لاختبارات الذكاء العام فيتناول الفصل السابع اختبارات الذكاء الفردية أما الفصل الثامن فسوف يتناول الاختبارات الجماعية للذكاء.

الفصل السابع

اختبارات الذكاء الفردية

يتناول هذا الفصل بالعرض اختبارات الذكاء الفردية، وعلى وجه الخصوص أشهر اختبارين فرديين للذكاء وهما: مقياس ستانفورد - بينيه، ومقياس وكسلر، ويعد مقياس بينيه على وجه الخصوص الوالد الشرعى لحركات قياس العقل الإنسانى - وخاصة عند الأطفال، ولذلك سيحظى فى هذا الفصل باهتمام خاص، ثم نعرض لمقاييس وكسلر، ولبعض الاختبارات والمقاييس الأخرى الأقل شهرة والتى لا تقل أهمية.

مقياس ستانفورد - بينيه

الفرد بينيه:

ألفرد بينيه (١٨٥٧ - ١٩١١) هو العالم الفذ الأصيل، بدأ تدريبه فى ميدان الطب ثم سرعان ما أصبح أشهر علماء النفس فى فرنسا فى عصره؛ وله فضله العظيم على علم النفس فى بلده، فقد أنشأ مع زميله بيونى Beauns أول معمل لعلم النفس عام ١٨٨٩ فى جامعة السوربون، وأول مجلة متخصصة باللغة الفرنسية «حولية علم النفس L'Année Psychologique عام ١٨٩٥.

ولقد بدأ بينيه عالما تجريبيا، فقبل أن يعد اختباراته العقلية المشهورة اهتم بدراسة العمليات العقلية المعرفية، وأطلق على منهجه اسم «سيكولوجية الفرد»، وفى رأيه أن أهم ما نجده من فوارق بين الناس إنما يكون هو فى الأنشطة العقلية العليا كالاستدلال والحكم، وهى جميعا من العمليات التى يصعب قياسها بالوحدات الفيزيائية (كالإحساس)، كما لا تتوفر فيها الوحدات الأولية للنشاط الذهنى، والتى كان يهتم بها أغلب علماء النفس من أنصار بنائية فوندت (ومنهم، تتشنر) (١٤).

وقد تأكدت صحة وجهة نظر بينيه من نتائج بعض الدراسات التى أجريت على صدق اختبارات جيمس ماكس كاتل (الذى اقتفى فيها أثر السير فرنسيس جالتون فى إنجلترا) لمقاييس عقلية، وبخاصة بحوث كلارك وزلر وستيلا شارب (٢٨)، ولذلك بدأ بينيه يطبق فى بحوثه عن الاختبارات العقلية ما توصل إليه من نتائج البحث

التجريبى فى عملية التفكير، فانتقد اختبارات كاتل وجالتون لأنها تقتصر على النواحى الحسية البسيطة، وفضل الاختبارات الأكثر تعقدا وتركيبا، واقترح عددا من الاختبارات لقياس «الملكات» العقلية (٦٣).

وقضى بينيه الفترة بين عامى ١٨٩٦ و ١٩٠٥ فى إجراء سلسلة طويلة من البحوث على الأطفال الأسوياء وغيرهم مطبقا أنواعا عديدة من الاختبارات لقياس مختلف «الملكات» العقلية بالإضافة إلى عدد من التجارب أجراها حول المسائل الخلافية فى سيكولوجية التفكير فى ذلك الوقت مثل التجريد، والتفكير بدون صور وزمن الرجع والتمييز الحسى، وكان الهدف الذى يسعى إليه من ذلك هو تحديد العلاقة بين هذه الاختبارات والعمر الزمنى والتحصيل المدرسى وتقديرات المعلمين للقدرة العقلية، وكانت بحوثه جيدة التصميم وأدت به إلى التحول من الاهتمام بالاختبارات التى تقيس «الملكات» التقليدية إلى تلك التى تقيس المستوى العقلى بالفعل.

وهكذا كانت نظرة بينيه إلى الذكاء أكثر شمولا مما ظهر فى مقياسه ذاته بعد ذلك؛ حين لجأ إلى استخدام درجة كلية واحدة هى العمر العقلى، فى قياس الذكاء، ويمكن للقارىء المهتم بإسهامات بينيه فى النظرية الأساسية للذكاء أن يرجع إلى (٢٨).

نشأة مقياس بينيه:

فى عام ١٩٠٤ طلبت وزارة المعارف الفرنسية من ألفرد بينيه والطبيب الفرنسى ثيوردور سيمون دراسة الطرق التى يمكن أن تستخدم فى تصنيف الأطفال الذين يتميزون ببطء التعلم فى المدارس الفرنسية، فعكفا على وضع مقياسهما المشهور للذكاء بغرض التمييز بين الأطفال الأسوياء وضعاف العقل.

ولم يكن الهدف دراسة وتحليل استعدادات المتخلفين عقليا، وإنما كان الهدف عمليا بحتا وهو تقويم وقياس ذكاء هؤلاء الأطفال بصفة عامة وتحديد مستواهم العقلى ومقارنتهم بالأطفال الأسوياء من نفس العمر والمستوى. أى الهدف كما يقول بينيه وسيمون (٦٣) هو بناء مقياس «مترى» Metrical للذكاء يتكون من عدد من «الاختبارات» المتدرجة فى الصعوبة، مرتبة حسب الأعمارحتى يستطيع الفاحص أن يحدد مدى تقدم الطفل عقليا أو تخلفه، وقد سعى بينيه إلى أن تكون هذه «الاختبارات»

متنوعة، مع التركيز على فئات معينة من العمليات العقلية اهتم بها دون غيرها عند إعداد المقياس وهى: الحكم والفهم العام والمبادأة والقدرة على التكيف.

ويتكون مقياس بينيه الأصلى الذى نشر عام ١٩٠٥ باسم مقياس بينيه ـ سيمون للذكاء Binet-Simon Scale من ٣٠ سؤالا أطلق على كل منها اسم «اختبار» وشاع هذا الاسم فى جميع الطبعات التالية لهذا المقياس فى جميع اللغات(*)، وفيما يلى قائمة بهذه الأسئلة، نذكرها كاملة لنبين إلى أى حد اختلفت عن اهتمامات علماء القياس العقلى قبل بينيه، وكيف أنها هيأت الموقف لظهور عدد كبير من الاختبارات العقلية تنحدر فى معظمها من صلب» مقياس بينيه على حد تعبير أنستازى (٦١) بالإضافة إلى شيوعها فى معظم تعديلات مقياس بينيه ذاته.

(١) التآزر البصرى (تتبع شمعة مضيئة بالرأس والعينين).
(٢) الفهم عن طريق ولمس (الإمساك بمكعب بعد لمسه).
(٣) الفهم عن طريق البصر أو الرؤية.
(٤) التعرف على الطعام (التمييز بين قطعة الخشب وقطعة الشيكولاتة).
(٥) البحث عن الطعام (الاستجابة لقطعة من الشيكولاته ملفوفة فى الورق).
(٦) تتبع التعليمات البسيطة أو الإيماءات المتكررة.
(٧) الإشارة إلى الأشياء وخاصة أعضاء الجسم (الرأس، الأنف... إلخ).
(٨) التعرف على الأشياء من صورها.
(١٠) التمييز الحسى بين طول خطين.
(١١) تكرار ثلاثة أرقام.
(١٢) التمييز بين وزن شيئين.
(١٣) القابلية للإيحاء المقارنة بين خطين متساويين (أو البحث عن أشياء لا وجود لها فى صورة).
(١٤) تعريف الكلمات البسيطة (التى تدل على أشياء مألوفة).
(١٥) تكرار جملة تتكون من ١٥ كلمة.

(*) آثرنا أن نشير إلى «الاختبارات» التى يتألف منها مقياس ستانفورد ـ بينيه بمصطلح «الأسئلة» فى متن هذا الكتاب، حتى لا تختلط بمعنى كلمة «الاختبار» كما تشيع فى التراث السيكولوجى الحديث.

(١٦) بيان أوجه الاختلاف بين شيئين.
(١٧) الذاكرة البصرية.
(١٨) رسم الأشكال من الذاكرة.
(١٩) مدى ذاكرة الأرقام.
(٢٠) بيان أوجه التشابه بين الأشياء.
(٢١) التمييز بين الخطوط بسرعة.
(٢٢) ترتيب ٥ أوزان.
(٢٣) تحديد الوزن الناقص (من بين الأوزان الخمسة فى الاختبار السابق).
(٢٤) إعطاء كلمات ذات قافية واحدة.
(٢٥) تكملة الجمل.
(٢٦) بناء جملة تحتوى على ٣ كلمات معينة.
(٢٧) الإجابة على الأسئلة (مثل: ماذا تفعل وأنت نعسان؟).
(٢٨) تحديد الوقت بعد أن تحل عقارب الساعة محل بعضها.
(٢٩) ثنى وتقطيع قطعة من الورق.
(٣٠) التمييز بين الكلمات المجردة (مثل: حزن وسأم).

وإذا فحصنا هذه القائمة فحصا دقيقا نجد أن الأسئلة الثلاثة الأولى هى فى الواقع أسئلة للنمو الحركى، أما الأسئلة الأخرى فيمكن أن تعد من النوع العقلى، ومن بينها ١٨ سؤالا (أى حوال الثلثين) من النوع الذى يقيس ما يسميه جيلفورد التفكير (المعرفة) و٥ أسئلة تقيس ما يسميه الذاكرة بأنماطها المختلفة (الأسئلة ١١، ١٥، ١٧، ١٨، ١٩) ولايوجد إلا ثلاثة أسئلة فقط تقيس ما أطلق عليه فيما بعد قدرات التفكير الابتكارى (أو الإنتاج التباعدى عند جيلفورد).

وكانت الطريقة التى استخدمها بينيه فى تقنين مقياسه غير ملائمة. فقد طبقه على ٥٠ طفلا (١٠ من كل فئة من فئات الأعمار ٣، ٥، ٧، ٩، ١١) اختارهم معلموهم على أساس أنهم متوسطو القدرة العقلية، بالإضافة إلى عدد من الأطفال لم يذكره بينيه كانوا من ضعاف العقول، وكانت طريقة التصحيح تعتمد على أساس تجريبى مبدئى يسمح بتصنيف الأطفال تبعا لأعلى مستوى يصل إليه الطفل؛ فالمعتوهون لايمكنهم تجاوز السؤال السادس، ولايتجاوز البلهاء السؤال الخامس عشر،

أما حدود المأفونين فلم تكن معروفة، رغم أنهم يعجزون فى العادة عن الإجابة على الأسئلة المجردة التى يستطيع معظم الأطفال العاديين من سن ١١ الإجابة عليها.

وبالطبع ارتبط عمر المفحوصين بنجاحهم فى الاختبار، ولم تكن لدى بينيه حتى ذلك الوقت (عام ١٩٠٥) طريقة تسمح بالمقارنة بين الأعمار، ولو أنه كان يلتمس طريقة فيما يبدو نحو مفهوم العمر العقلى، يقول بينيه وسيمون عام ١٩٠٥ (٦٣ : ٣٢١).

«ومن المحتمل أن ننجح فى يوم ما فى إيجاد علامات على التخلف السيكولوجى مستقلة تماما عن العمر، وقد يكون من المفيد جدا التعرف على هذه العلامات ، إلا أننا فى الوقت الحاضر نجد أن ما يجذب انتباهنا أكثر من غيره هو التشابه بين الأسوياء من صغار السن وغير الأسوياء من كبار السن، وهذا التشابه من الكثرة والغرابة بحيث أننا لو حصلنا على استجابات طفل لا نعرف عمره لا نستطيع الحكم عليه بالسواء أو عدم السواء.

وكانت لطبعة عام ١٩٠٥ من مقياس بينيه عدة سمات تعد نقطة تحول فى تاريخ القياس العقلى ففيه نجد لأول مرة أن الأسئلة التى تقيس مكونات عديدة ترتبط فيما بينها لتحدد مستوى من الأداء بدلا من القياس المنفصل لأبعاد التحليل السيكولوجى التقليدى الذى كان شائعا فى عصر بينيه مثل الملكات أو عتبات الإحساس أو غيرها وبالإضافة إلى ذلك فإن وضع سؤال فى المقياس أو حذفه أو تعديله كان من المسائل التجريبية ، ويدل على خبرة بينيه الطويلة بالتجريب العملى، فقد كان يهتم بمدى العلاقة بين السؤال وأحكام المحك (أى بنوع من الصدق) ، ثم إن هذه الاختبارات لم تكن محددات معملية دقيقة وإنما كانت أدوات عملية بسيطة ومهمة لدراسة الأطفال فى مواقف واقعية (مثل حجرة الدراسة).

وقد أدى ظهور مقياس ١٩٠٥ إلى جذب انتباه عدد كبير من المهتمين بمشكلات الضعف العقلى والتصنيف التربوى، فقام العالم الأمريكى جودارد Goddaard بترجمته إلى الإنجليزية، كما جربه ديكرولى Decroly وديجان Degand فى بلجيكا، وكان الكثيرون - ومنهم بيرت فى إنجلترا وترمان فى أمريكا - على اتصال مباشر ببينيه كما اقترحوا فى الوقت نفسه اختباراتهم الخاصة.

تعديل عام ١٩٠٨ :

فى عام ١٩٠٨ ظهر مقياس جديد لبينيه وسيمون يعتمد اعتمادا كبيرا على المقياس السابق ومع تعديلات جوهرية، فقد أصبح اسم المقياس الجديد «نمو الذكاء عند الأطفال The development of intelligence in children وتحول تركيز هذا المقياس من غير الأسوياء إلى الأسواء، واستبعدت اختبارات المعتوهين، واحتفظ المقياس بخاصية استخدام الأسئلة القصيرة، العديدة المتنوعة من حيث المحتوى، والمرتبة ترتيبا تجريبيا حسب الصعوبة، ولأول مرة تصنف الأسئلة حسب مستويات الأعمار ابتداء من سن ٣ سنوات حتى سن ١٣ سنة تبعا للعمر الذى يستطيع الأطفال الأسوياء الوصول إلى أسئلته بنجاح، وكان تحديد الأسئلة للمستويات العمرية المختلفة تحديدا جزافيا، ليس بسبب قلة عدد أفراد عينة التقنين فحسب وإنما أيضاً بسبب الصعوبات العملية فى إعداد الأسئلة التى تقع فى مستويات الصعوبة المطلوبة ، وبوجه عام اعتبر المستوى العمرى للاختبار هو ما يستطيع ما بين ٥٥ ٪، ٧٥ ٪ من أطفال التقنين بالإجابة عليه إجابة صحيحة، لأن هذا المدى يتضمن نسبة الـ ٥٠ ٪ التى تدل دلالة عامة على الأداء المتوسط وكانت عينة التقنين ٢٠٣ أطفال تتراوح أعمارهم بين ٣ سنوات و١٣ سنة.

وقد أدى ترتيب الأسئلة النمائى حسب مستويات العمر بالفاحصين إلى التعبير عن المستوى الارتقائى للطفل بالعمر الذى يحصل فيه الطفل المتوسط على درجة مكافئة، ورغم أنه لم تتحدد مستويات النجاح والفشل على أساس أسئلة المقياس، فإن العمر العقلى للطفل كان يتحدد على أساس يساوى العمر الذى يستطيع أن يجتازه ويجتاز الأعمار السابقة عليه، (العمر القاعدى أو الأساسى) ثم يضاف إليه عام إضافى من العمر العقلى لكل ٥ أسئلة من المستويات الأكثر صعوبة.

هذه هى الخصائص الأساسية لمعيار العمر العقلى كما اقترحه تعديل ١٩٠٨ ، وهو المعيار الذى اقتبسته مقاييس الذكاء اللاحقة، ورغم أن هذا المعيار يتضمن مشكلات إحصائية أشرنا إليها فى الفصل السادس، فإن له فائدته فى أنه يوفر للباحث طريقة سهلة الفهم فى وصف الفروق بين الأطفال فى المستوى العام للقدرة العقلية، ومع ذلك بقيت مشكلتان هامتان لم يحلهما مقياس ١٩٠٨ : أولاهما أن عدد الأسئلة فى كل

مستوى عمرى يختلف، ويمتد بين ٣، ٨ أسئلة،وثانيهما أن درجة العمر العقلى لاتفهم كمعيار للتفوق أو التخلف إلا إذا ارتبطت بالعمر الزمنى للطفل.

تعديل عام ١٩١١:

فى عام ١٩١١ قام بينيه - قبيل وفاته المبكرة - بتعديل مقياسه ونشره باسمه منفردا، فأعاد ترتيب كثير من الأسئلة، ووحد عددها فى كل مستوى عمرى فجعله خمسة أسئلة ثم أضاف عددا آخر من الأسئلة (٥ لمستوى ١٥ سنة، و٥ لمستوى الراشد) وهكذا أمكن - بعد تساوى عدد الاختبارات فى كل مستوى - إعطاء كسور من العمر العقلى مقدارها سنة لكل اختبار يستطيع المفحوص الإجابة عليه بعد عمره القاعدى أوالأساسى، ومع ذلك فإن مفهوم العمر العقلى ظل على حاله، أى لايمكن استخدامه كمؤشر على المستوى العقلى مستقلا عن العمر الزمنى، وقد أدى ذلك بالعالم الألمانى شترن Stern عام ١٩١٢ إلى اقتراح معيار النسبة العقلية أى $\frac{\text{العمر العقلى}}{\text{العمر الزمنى}}$ وهو المعيار الذى استخدمه - بعد التعديل - ترمان كما سنوضح بعد قليل. ويوضح الجدول (١٧) الأسئلة التى يتكون منها مقياس عام ١٩١١، ومن هذا الجدول يتضح أن المقياس فى صورته الأخيرة على يد بينيه يتكون من ٥٤ سؤالا، وفيه حذفت المستويات الخاصة بالأعمار ١١، ١٣، ١٥ سنة كما أن أسئلة سن ٤ كانت ٤ أسئلة فقط (٦٦). راجع جدول (١٧).

جدول رقم (١٧)

يبين الاختبارات التى يتكون منها مقياس بينيه (تعديل عام ١٩١١)

العمر	أسئلة تعديل عام ١٩١١				
٣	اسم العائلة	الإشارة	عددان	٦ مقاطع	صورة (سهل)
٤	الجنس	التسمية	٣ أعداد	خطان	ـــــ
٥	٤ قطع نقود	المربع	وزنان	١٠ مقاطع	بطاقة مقسمة
٦	١٣ قطعة نقود	المعين	مقارنة الوجوه	تعريف الكلمات (سهل)	الصباح والمساء
٧	٣ قطع نقود، ٣ أنصاف لهذه القطع	٤ ألوان	اليمين والشمال	الترتيب الثلاثى	صورة (صعب)
٨	السمات الناقصة	التاريخ	٥ أعداد	الفروق (المحسوسة)	العد من ٢٠ إلى١
٩	٩ قطع نقود	التفسير	الشهور	تعريف الكلمات (محسوس)	الأسئلة السهلة
١٠	بناء الجمل (سهل)	الرسم من الذاكرة	٥ أوزان	السخافات	الأسئلة الصعبة
١٢	بناء الجمل (صعب)	الإيحاء	٦٠ كلمة فى ٣ دقائق	تعريف الكلمات (مجرد)	الجمل المختلطة
١٥	المشكلات والمسائل	٣ إيقاعات	٧ أعداد	٢٦ مقطعا	الصورة
الرشد	الورق المطوى	إعادة التعبير	الفرق بين حرفين	الفروق المجردة	المثلث المعكوس

مقياس ستانفورد ـ بينيه لترمان:

لم تكن محاولة جو دارد ترجمة ونقل مقياس بينيه إلى اللغة الإنجليزية هى المحاولة الوحيدة، ففى عام ١٩١٠ نشر هوى Huey وويل Ehipple ترجمة له، كما نشر فالين Wallin وكولمان Kuhlmann ترجمة أخرى فى عام ١٩١١. وفى عام ١٩١٣ بدأ بيرت Burt نقل المقياس إلى الإنجليزية وتقنينه على عينة كبيرة من الأطفال الإنجليز عندما عين أخصائيا نفسيا لمقاطعة لندن، وظل يتابع مشروعه هذا طوال تاريخه العلمى كأستاذ لعلم النفس بجامعة لندن، بالإضافة إلى نقل المقياس فى أقطار عديدة بلغاتها مثل إيطاليا وألمانيا وروسيا والسويد وتركيا واليابان، وفى عام ١٩٢٠ ترجم أحمد فكرى (٥) مقياس بينيه الأصلى إلى اللغة العربية، وقدم له تعليمات مقتضبة، وكان يعتقد خطأ أن المدرس ـ دون تدريب سابق ـ يمكنه تطبيق المقياس والحكم على ذكاء الطفل. ومع ذلك فإن أكثر هذه المحاولات شهرة فى تاريخ القياس النفسى محاولة لويس ترمان L. M. Terman (١٨٧٧ – ١٩٥٦) أستاذ علم النفس بجامعة ستانفورد،والتى عرف بها المقياس باسم مقياس ستانفورد ـ بينيه للذكاء، وقد شاعت هذه التسمية أكثر من الاسم الأصلى للمقياس.

وقد ظهرت الطبعة الأولى لمقياس ستانفورد ـ بينيه عام ١٩١٦. وأدخل ترمان على المقياس الأصلى عددا من التعديلات والإضافات الجوهرية بحيث يمكن القول أنه اختبار جديد، فقد كان حوالى ثلث الأسئلة جديدا، وعدل ترمان عددا من الأسئلة القديمة وأعيد توزيعها على المستويات العمرية المختلفة أو حذفت، وأعيد تقنين المقياس كله على عينة من الأمريكيين بلغ عددها ١٠٠٠ طفل و٤٠٠ راشد، وأعدت تعليمات مفصلة لتطبيق كل سؤال وتصحيحه، واستخدام معيار نسبة الذكاء (تعديلا لمعيار النسبة العقلية لشترن وكولمان) فى صورتها التى شاعت فى القياس النفسى هى:

$$\text{نسبة الذكاء} = \frac{\text{العمر العقلى}}{\text{العمر الزمنى}} \times ١٠٠$$

وقد كتبت مجلة المقتطف المصرية مقالا عن هذا المقياس عام ١٩٢٧ (٥٥) كما كتب عنه أمير بقطر عام ١٩٢٨ (١٣). ونقله إلى اللغة العربية لأول مرة طبيب غير

متخصص فى علم النفس هو حسن عمر (٢٢) . ثم ترجمه على نحو أدق إسماعيل القبانى (٦) عام ١٩٣٨ .

وفى عام ١٩٣٧ ظهرت فى الولايات المتحدة الطبعة الثانية المعدلة للمقياس، وتتكون من صورتين متكافئتين، الصورة ل، والصورة م. وفى هذه الطبعة زاد عدد أسئلة المقياس وأعيد تقنينه تماما على عينة جديدة جيدة السحب من المجتمع الأمريكى بلغ عددها أكثر من ٣٠٠٠ مفحوص تتراوح أعمارهم بين سن ١,٥ سنة إلى سن ١٨ ، واشتملت كل عينة عمرية على عدد متساو من الذكور والإناث، وتم اختبار معظم المفحوصين من سن ٦ سنوات وما بعدها من تلاميذ المدارس، أما عينات أطفال ما قبل المدرسة (عينات الأعمار ١,٥ ، ٢ ، ٢,٥ ، ٣ ، ٣,٥ ، ٤ ، ٤,٥ ، ٥ ، ٥,٥ سنة) فتم الحصول عليها بطرق مختلفة أيسرها تطبيق المقياس على أخوة أطفال وتلاميذ المدارس، وقد روعى فى اشتقاق العينات التوزيع الجغرافى والمستويات الاقتصادية والاجتماعية.

وقد ظهرت الطبعة التجريبية للصورة ل باللغة العربية ١٩٥٦ وقام بإعدادها محمد عبد السلام أحمد ولويس كامل مليكه (٤٧) .

وفى عام ١٩٦٠ ظهرت طبعة ثالثة لمقياس ستانفورد ـ بينيه فى صورة واحدة فقط تجمع أفضل الأسئلة من صورتى طبعة عام ١٩٣٧ ، وبدون تقديم أى محتوى جديد أو صياغة أسئلة جديدة كان من الممكن حذف الأسئلة التى لا تساير العصر، وأعيد توزيع الأسئلة التى تغيرت مستويات صعوبتها فى الفترة بين التعديلين نتيجة للتغيرات الثقافية، وفى إعداد طبعة عام ١٩٦٠ واجه ترمان مشكلة شائعة فى القياس النفسى وتتلخص (٦١) فى أنه من المستحب عامة إدخال تعديلات دورية على المقاييس للإفادة من التقدم العلمى فى ميدان بناء الاختبارات، ومن الخبرة السابقة باستخدام الاختبار، بالإضافة إلى تحديث محتوى الاختبار على نحو يجعله مسايرا لروح العصر، وتزداد الحاجة إلى الاعتبار الأخير بالنسبة لأسئلة المعلومات والمواد المصورة والتى تتغير تبعا لتغير الظروف الثقافية، فاستخدام أسئلة قديمة يؤثر فى الصدق الظاهرى للاختبار دون شك، كما قد يغير فى مستوى صعوبة الأسئلة، ومن ناحية أخرى فإن التعديل قد يجعل معظم البيانات المتوافرة من البحوث السابقة على

الاختبار لا ترتبط بالصورة الجديدة، فالاختبارات التى استخدمت استخدامات واسعة لمدة سنوات كمقياس ستانفورد ـ بينيه، تتوافر عنه ثروة من المعلومات يجب أن توزن فى مقابل الحاجة إلى التعديل، ولهذه الأسباب فضل تلاميذ ترمان الجمع بين الصورتين فى صورة واحدة بعد انتقاء أفضل ما فيهما، وفقدان صورة مكافئة ليس ثمنا فادحا لتحقيق هذه الغاية، وخاصة أنه أصبحت توجد منذ عام ١٩٦٠ اختبارات فردية أخرى جيدة الصنع لم تكن متوافرة عام ١٩٣٧ .

وفى عام ١٩٧٢ أجرى تقنين جديد للصورة م ـ ل بإشراف روبرت ثورنديك (٨٠) دون إدخال تعديلات تذكر على محتوى الاختبار، أما التعديل الجوهرى فقد طرأ على المعايير التى اشتقت من أداء عينة جديدة بلغت ٢١٠٠ مفحوصا تم اختبارهم خلال العام الدراسى ١٩٧١ – ١٩٧٢، وذلك بمعدل ١٠٠ مفحوصا لكل مستوى عمرى (٢–١٨ سنة) كما استخدمت نسبة الذكاء الإنحرافية بمتوسط مقداره ١٠٠ وانحراف معيارى ١٦ . وقد لوحظ فى هذا التقنين ارتفاع مستوى الأداء وبصفة خاصة الأعمار الصغيرة (اناستازى، أوريانا ١٩٩٧ :٢٠٦) وكان هذا التحسن بمقدار ١٥ وحدة ذكاء (IQ Points) ، وأرجع هذا التحسن إلى أثر وسائل الإعلام على الأطفال ونقصان معدل الأمية فى السبعينيات والثمانينيات وزيادة من يستمرون فى التعليم. وفى تحليل قام به ثورنديك عام ١٩٧٧ أرجع هذه الزيادة إلى البرامج التليفزيونية التى تستثير النمو المعرفى للاطفال وحتى تكون لهذه العينة خصائص التمثيل رغم الصعوبة العملية فى تطبيق الاختبارات الفردية لجأ ناشرو الاختبار إلى عينة أكبر بلغت مائتى ألف تلميذ وطالب من الصفوف الثالث حتى الثانى عشر (فى النظام التعليمى الأمريكى) استخدمت فى تقنين أحد الاختبارات الجماعية المشهورة (اختبار القدرات المعرفية الذى سنشير إليه فى الفصل الثامن) ومثلت خصائص المجتمع الأمريكى من حيث الحجم والمناطق الجغرافية والمستويات الاقتصادية الاجتماعية والأصول العنصرية.

وفى اختيار عينة تقنين مقياس ستانفورد ـ بينيه من بين هذه العينة الأكبر اعتمد على درجات المفحوصين فى البطارية اللفظية فى اختبار القدرات المعرفية المشار إليه بحيث يصبح توزيع درجات العينة الجديدة مطابقا للتوزيع العام على مستوى العينة الكلية، ولم يستبعد إلا أولئك الذين لايتحدثون اللغة الإنجليزية، وحتى تشمل العينة

الأعمار من ٢ – ٨ سنوات اختار الباحثون أخوة المفحوصين الذين تضمنتهم العينة وبالمثل اختيرت بعض الحالات الإضافية فى الأعمار العليا، وهكذا أصبحت عينة تقنين المقياس مؤلفة من مائة مفحوص تقريبا لكل فئة عمرية مداها نصف سنة للأعمار من ٢– $\frac{١}{٢}$ ٥، ومائة مفحوص تقريبا أيضًا لكل فئة عمرية مداها سنة للأعمار من ٦ – ١٨ سنة.

وفى عام ١٩٨٥ صدرت الطبعة الرابعة من المقياس من إعداد روبرت ثورنديك واليزابيث هاجن وجيروم ساتلر (٨١)، وقد حافظت هذه الطبعة الجديدة على المعالم التاريخية للمقياس التى استمرت فى الطبعات السابقة، وبصفة خاصة الملاحظات الكلينيكية. فأبقيت الأسئلة الصالحة من طبعتى ١٩٣٧، ١٩٦٠، وأضيفت أسئلة أخرى جعلته مقياسا حديثاً من حيث المحتوى والخصائص الإحصائية وإجراءات التطبيق، واستخدمت فى المقياس الجديد أحدث النماذج السيكومترية وخاصة نظرية السمات الكامنة (أو نظرية الاستجابة على المفردة item response theory ، وحلل تحيز الأسئلة سواء من حيث الصياغة أو من حيث الخصائص الإحصائية.

ولعل أهم التعديلات التى أدخلت على المقياس فى هذه الطبعة استخدام مفاهيم القياس التكيفى للقدرة adaptive ability وذلك بتطبيق نظام الدرجات المتعددة وليس الدرجة الواحدة كما كان الحال فى هذا المقياس طوال تاريخه وأمكن بهذه الطريقة إعطاء المفحوص أربعة أنواع من الدرجات خلال زمن الاختبار هى: الاستدلال اللفظى، والاستدلال الكمى، والاستدلال المجرد فى مقابل الاستدلال البصرى، والذاكرة قصيرة المدى، وبالإضافة إلى ذلك يمكن أن يعطى المفحوص درجة كلية مركبة فى المقياس.

وعلى الرغم من أن كل درجة من هذه الدرجات تعكس جزئيا أحد عوامل القدرة العقلية العامة المشتركة فى جميع القدرات، وعلى الرغم أيضًا من الدرجة الكلية المركبة من هذه الدرجات الأربع تعطى تقديرا ثابتا للعامل العام، فإن كل درجة منها لها معناها السيكولوجى الخاص والذى له أهميته فى ذاته، وقد حسبت المعايير لكل درجة من الدرجات الأربع والدرجة الكلية، وقد نظم المقياس تبعا لمستويات تبدأ من الطفل الأقل من المتوسط من سن سنتين وحتى مستوى الراشد المتفوق.

وتتضمن مفردات المقياس أدوات مختلفة تبنى عليها المفردات مثل المكعبات واللوحات المختلفة الألوان والأشكال، ولوحة كبيرة لصورة عروسة وكتيب لتسجيل استجابات المفحوص، وكتاب مرشد للفاحص لطبيعة التطبيق، ويجدر القول أن تطبيق المقياس يتطلب من الفاحص تدريبا واعدادا جيدا للتطبيق والتصحيح وتفسير الدرجات، ومعرفة جيدة بخطوات السير فى تطبيق المقياس، كما أن التعليمات التى يجب أن يستخدمها الفاحص ملزمة عند التطبيق والتصحيح المتزامن للتطبيق. ويتميز التطبيق الفردى بإمكانية ملاحظة أداء المفحوص وطريقته فى معالجة المشكلات التى يتضمنها المقياس وكذلك ملاحظة انفعالات المفحوص أثناء التطبيق وقدرته على التركيز وثقته بنفسه.

وهذه الملاحظات الكيفية يجب وضعها فى الإعتبار وتفسيرها بطريقة موضوعية، وتعتمد موضوعية التفسير على مهارة الفاحص فى الملاحظة والتسجيل.

أما من حيث مفردات المقياس المعدل فإن كل فئة من المفردات تتجمع فى اختبار واحد مستقل يتدرج من حيث الصعوبة، ويوجد بهذه الطبعة خمسة عشرة اختبارا تقيس الجوانب المعرفية الاربعة السابق الاشارة لها، ويتم تطبيق هذه الإختبارات فى متتالية مختلطة بحيث تستثير المفحوص وتزيد من اهتمامه.

ويغطى مستوى الصعوبة فى ست اختبارات المدى الكلى للعمر الزمنى الذى يغطيه الإختبار، أما بالنسبة للتسع اختبارات الأخرى فتبدأ إما على مستوى صعب أو تنتهى عند مستوى سهل وذلك طبقا لطبيعة المفردات (أنظر الشكل – عن انستازى وأوربانا ١٩٩٧ : ٢٠٩).

شكل (٢)، مجموعات الإختبارات الفرعية ومداها (تعديل ١٩٨٤)

ملحوظة (فى الشكل ٢) :

الإختبارات التسعة المرسومة هكذا (xxx) والمحددة فى أعمار معينة استخدمت لتقدير العمر العقلى لعينات كان أدائها إما أقل أو أعلى بطريقة ملحوظة أثناء تقنين الاختبار، وقد استخدمت درجاتهم لتقدير العمر العقلى لهم وأدرجت فى جداول المعايير (إلا أنها تستخدم بحذر) ويمكن الرجوع لدليل الإختبار الأصلى لمزيد من التفصيل.

(عن أناستازى وأوربانا (بتصريف) ١٩٩٧ : ٢٠٩)

ويتطلب تطبيق طبعه ١٩٨٤ مرحلتان. ففى المرحلة الأولى يبدأ الفاحص من اختبار الكلمات بإعتباره محورا لبدء الإختبارات الأخرى، ونظرا لأن اختبار الكلمات يغطى العمر الزمنى من ٢-١٨ سنة فإن البدء فيه يعتمد على العمر الزمنى للمفحوص، أما باقى الإختبارات فيعتمد البدء فيها على درجة المفحوص فى اختبار الكلمات وعمره الزمنى، وهذه الحدود مبينة فى جدول فى كراسة التعليمات الخاصة بالاختبار.

وفى المرحلة الثانية من تطبيق الاختبار يقوم الفاحص بتحديد العمر القاعدى والسقف (Ceiling) لكل اختبار وذلك بناء على الاداء الفعلى للمفحوص. ويتحدد مستوى العمر القاعدى عندما يجتاز المفحوص أربعة مفردات من مستويين متتاليين، وإذا لم يتحدد هذا المستوى من البداية فعلى الفاحص الإنتقال إلى مستويات أدنى. أما سقف الاختبار فيتحدد عندما يفشل المفحوص فى ثلاثة أو الأربعة مفردات جميعهم فى مستويين متتاليين وتتحدد الدرجة للمفحوص طبقا للأداء الفعلى له ومقارنتها بما جاء فى دليل الاختبار.

ويتطلب تطبيق المقياس فترة زمنية متصلة تتراوح بين ٣٠-٩٠ دقيقة علما بأنه توجد بعض الأمثلة فى ١١ اختبارا لا تضاف درجتها إلى درجة المفحوص فى الاختبار، كما يوجد بعض المفردات التى تتطلب استجابات حرة فى اختبارات الكلمات والفهم والسخافات، والنسخ والعلاقات اللفظية.

وتتحدد معايير الاختبار لكل فئة من الفئات الاربع للإختبار بدرجة معيارية متوسطها ٥٠ وانحرافها المعيارى ٨، كما توجد معايير للدرجة الكلية متوسطها ١٠٠ وانحرافها المعيارى ١٦ ويطلق عليها Standard Age Score (SAS)

وعلى الرغم من أن كل درجة من هذه الدرجات تعكس جزئيا أحد عوامل القدرة العقلية العامة المشتركة فى جميع القدرات، وعلى الرغم أيضا من الدرجة الكلية المركبة من هذه الدرجات الأربع تعطى تقديرا ثابتا للعامل العام، فإن كل درجة منها لها معناها السيكولوجى الخاص والذى له أهمية فى ذاته، وقد حسبت المعايير لكل

درجة والدرجة الكلية، وقد نظم المقياس لمستويات تبدأ من الطفل الأقل من المتوسط من سن سنتين وحتى مستوى الراشد المتفوق.

تحليل أسئلة المقياس وتطويرها:

ذكرنا فى الصفحات السابقة نشأة وتطور مقياس بينية والتعديلات التى أدخلت على هذا المقياس الهام، والجهود التى بذلت فى تحديثه وترجمته فى مجتمعات مختلفة، فطبعه عام ١٩٣٧ لمقياس ستانفورد ـ بينيه أكثر من عشر سنوات من البحث العلمى المنظم حول تجميع الأسئلة الجيدة وتجريبها على جماعات صغيرة العدد من المفحوصين الذين طبق عليهم مقياس عام ١٩١٦ . وكانت الخطوة الأساسية فى ذلك هى تطبيق الصورتين المبدئيتين فى ذلك الوقت على عينة التقنين التى أشرنا إليها (٣١٨٤ مفحوصا) ومنها تم الانتقاء النهائى للأسئلة وتوزيعها على الأعمار وقد أعطيت لكل مفحوص صورتا الاختبار بفاصل زمنى يتراوح بين يوم واحد وأسبوع واحد، وفى هذه الخطوة كان محك تحليل الأسئلة لتحديد صدقها هو العمر الزمنى والدرجة الكلية المركبة من الصورتين المبدئيتين، ويرتبط المحك الثانى بما أسميناه الاتساق الداخلى حتى يتحقق فى المقياس درجة من التجانس. واستخدمت ثلاثة مقاييس فى تحليل كل سؤال فى المقياس هى:

١ – منحنيات النسب المئوية للمفحوصين الذين يجتازون السؤال بنجاح فى الأعمار الزمنية المتتابعة.

٢ – منحنيات النسب المئوية للمفحوصين الذين يجتازون السؤال بنجاح فى الفئات المتتابعة للدرجة الكلية فى المقياس (وبالطبع لكل صورة من صورتى المقياس منحنياتها الخاصة).

٣ – معامل الارتباط بين كل سؤال والدرجة الكلية فى المقياس.

وتتفق الطريقة الأولى مع أسلوب بينيه الذى كان يستخدمه فى تحليل المفردات وانتقاء الأسئلة التى تصلح للمستويات العمرية المختلفة، وفى هذا يجب أن نشير إلى أن ترمان استخدم للحكم على صلاحية الأسئلة نسبا مئوية فى الأعمار الصغيرة أعلى

منها فى الأعمار الكبيرة. ومن ذلك مثلا أنه كان يعتبر السؤال مناسبا لسن ٣ سنوات إذا أجاب عليه ٧٣٪ من أطفال هذا السن إجابة صحيحة، أما فى حالة الأطفال من سن ١٠ سنوات فكانت هذه النسبة ٥٩٪، ويشير ماكنمر (٨٨) إلى أن هذا الاختلاف فى النسب المئوية هو من المطالب الضرورية فى مقاييس العمر، فمن الملاحظ أن الانحراف المعيارى للعمر العقلى يجب أن يتزايد مع العمر إذا كان لابد لنسبة الذكاء أن تظل ثابتة، وعلى ذلك فإذا كان التباين فى العمر العقلى فى الأعمار العليا أكبر منه فى الأعمار الدنيا فإن ذلك يعنى ضرورة وجود انتشار أكبر فى أداء المفحوصين الكبار فى مستويات الأعمار المتجاورة، وعلى ذلك فمع الزيادة فى العمر يقل عدد المفحوصين الذين يجتازون بنجاح السؤال المطابق لمستواهم العمرى.

ويمكن أن نوضح هذه المسألة بمثال بسيط لنفرض أنه بين الأطفال من سن ٥ سنوات يصل ٤٠٪ منهم إلى العمر العقلى ٥، بينما يصل ٣٠٪ إلى أعمار أعلى، و٣٠٪ أخرون إلى أعمار أقل من هذا المستوى، وحيث أن أى مفحوص عمره العقلى ٥ سنوات أو أكبر يجب أن يجتاز من الوجهة النظرية السؤال الخاص بسن ٥ سنوات فإن ٧٠٪ (٤٠٪ + ٣٠٪) من الأطفال من فئة سن ٥ سنوات نتوقع نجاحهم فى أسئلة سن ٥ سنوات، ولنفرض من ناحية أخرى أنه من بين الأطفال من سن ١٠ سنوات لا يصل إلى العمر العقلى ١٠ سنوات إلا ٢٠٪ بينما يصل إلى أعمار أعلى من ذلك ٤٠٪ منهم، وإلى أعمار أقل من ذلك ٤٠٪ آخرون، وبالطبع لايمكن الحصول على مثل هذا التوزيع إلا إذا كان التباين فى الأعمار العقلية لأطفال سن ١٠ سنوات أكبر منه فى سن ٥ سنوات، ويعنى ذلك أنه من بين الأطفال من سن ١٠ سنوات تصبح النسبة المتوقعة لاجتياز الأسئلة المخصصة لهذا السن هى (٢٠٪ + ٤٠٪ = ٦٠٪) وفى هذا المثال تصبح النسبة المئوية التى يمكن استخدامها محكا لانتقاء الأسئلة فى سن ٥ سنوات هى ٧٠٪، وفى سن ١٠ سنوات هى ٦٠٪.

وفى الممارسة الواقعية لانتقاء أسئلة مقياس ستانفورد - بينيه نجد أن المحك يتناقص من ٧٧٪ عند مستوى سن عامين إلى ما هو أقل قليلا من ٥٠٪ فى مستوى الرشد العادى، وتقل النسب عن هذا فى انتقاء أسئلة مستويات الرشد المتفوق وذلك

للوصول إل سقف مناسب للاختبار، وعند الانتقاء النهائى للمفردات لايقتصر الاهتمام على تمايز العمر والاتساق الداخلى فحسب، وإنما امتد ليشمل أيضًا مدى تناقض الفروق بين الجنسين وتوازنها فى النسب المئوية للناجحين فى مختلف المستويات العمرية، واستبعدت بالفعل المفردات التى اختلفت فيها أداء أحد الجنسين عن الآخر بنسب لها دلالة إحصائية، على أساس أن مثل هذه المفردات قد تدل على تباين اصطناعى أو غير مرتبط بما يقيسه الاختبار نتيجة للخبرات السابقة للجنسين.

وفى اعداد طبعة عام ١٩٦٠ من المقياس اختيرت الأسئلة من الصورتين ل. م على أساس أداء ٤٤٩٨ مفحوصا يمتد أعمارهم بين ٢,٥ سنة ، ١٨ سنة أجابوا على إحدى صورتى المقياس أو كليهما فى الفترة من ١٩٥٠ وحتى ١٩٥٤، كما تضمنت العينة مجموعتين طبقيتين اشتملتا على ١٠٠ طفل من سن ٦ سنوات صنفت حسب مهنة الوالد ، ١٠٠ مراهق من سن ١٥ سنة صنفت حسب مهنة الوالد وتوزيع المستوى التعليمى.

ويعنى هذا أن طبعة ١٩٦٠ لم تتضمن إعادة لتقنين المقياس أو إعادة لبناء معاييره، واستخدمت العينات الجديدة لمراجعة التغيرات التى طرأت على صعوبة الأسئلة فقط بعد ما تحددت مستويات الصعوبة بإيجاد النسب المئوية للأطفال الذين يجتازون السؤال فى أعمار عقلية متتابعة فى صورتى عام ١٩٣٧ وهكذا فإنه بالنسبة لتحليل المفردات تم تجميع المفحوصين تبعا للعمر العقلى الذى تحدده طبعة عام ١٩٣٧ وليس تبعا للعمر الزمنى وعلى هذا فإن الأعمار العقلية ونسب الذكاء فى طبعة ١٩٦٠ ظلت على النحو الذى كانت عليه فى طبعة ١٩٣٧ وفى ضوء أداء عينات تقنين هذه الطبعة. صحيح أنه روجعت بعض الأسئلة للتحقق من احتمال وجود فروق جغرافية و اقتصادية اجتماعية وذلك من خلال المقارنة بين المجموعات الفرعية التى تكون العينة الكلية، إلا أن ما يجب أن ننبه إليه أن المقياس لم يتضمن أى مواد جديدة، وإنما عدلت بعض الرسوم الخاصة بالأشياء الشائعة مسايرة للعصر، وفى هذه الأحوال اختيرت الأسئلة المعدلة قبليا على عينات صغيرة قبل تضمينها فى المقياس، وقد ظلت هذه الاستراتيجية فى انتقاء المفردات ثابتة فى طبعتى ١٩٧٢، ١٩٨٥.

وصف المقياس:

يتألف مقياس ستانفورد - بينيه من صندوق به مجموعات من اللعب المقننة

تستخدم مع الأعمار الدنيا، وكتيبين من البطاقات المطبوعة ، وكراسة لتسجيل الإجابات، وكراسة التعليمات، وكراسة معايير التصحيح.

وقد رتبت أسئلة المقياس وجمعت تبعا لمستويات الأعمار التى امتد من سن عامين حتى سن الرشد المتفوق، وتتحدد مستويات المقياس من سن ٢ حتى سن ٥ سنوات في فئات نصف سنوية على النحو الآتى: ٢، ٢ – ٦، ٣، ٣ – ٦، ٤ – ٦، ٥ (مع ملاحظة أن الرمز ٢ – ٦، ٣ – ٦، ٤ – ٦ يدل على سن عامين و٦ شهور،و٣ سنوات و٦ شهور، ٤ سنوات و٦ شهور على التوالى) وذلك بسبب التقدم العقلى السريع فى الأعمار الدنيا، أما الفترات من سن ٥ سنوات حتى ١٤ سنة فتتحدد مستويات الأعمار فيها فى فئات سنوية، وتدل المستويات الأخرى فى المقياس على الراشد المتوسط، والراشد المتفوق (١، ٢، ٣)، وكان عدد أسئلة كل مستوى عمرى ٦ ، باستثناء مستوى الراشد المتوسط الذى يبلغ عدد أسئلته ٨.

وكانت الأسئلة فى كل مستوى عمرى من مستوى صعوبة موحدة تقريبا، وترتب بصرف النظر عن أى فروق طفيفة فى الصعوبة، وقد أعد سؤال احتياطى فى كل مستوى عمرى معادل فى صعوبته لأسئلة العمر، ويستخدم هذا السؤال الاحتياطى إذا تطلب الأمر استبعاد أحد الأسئلة العادية بسبب بعض الظروف الخاصة تجعله غير ملائم للمفحوص، أو بسبب عدم الالتزام فى التطبيق بحرفية التعليمات بما يتعارض مع تقنين المقياس وهكذا يبلغ العدد الكلى لاختبارات المقياس ١٤٢ اختبارا (١٢٢ اختبارا أصليا، ٢٠ اختبارا احتياطيا).

ومعنى هذا أنه ابتداء من طبعة ١٩٦٠ من المقياس نجد سؤالا احتياطيا لكل مستوى عمرى، وهذا ما لم يتوافر فى طبعة ١٩٣٧ التى نقلت إلى اللغة العربية حيث اقتصر وجود الأسئلة الاحتياطية على المستويات العمرية من سن ٢ حتى سن ٥ سنوات فقط.

وقد اختير ٤ أسئلة من كل مستوى عمرى على أساس الصدق والتمثيل ليتألف منها المقياس المختصر الذى يستخدم حين لا يسمح الوقت بإعطاء المقياس الكلى، وقد وضعت أمام هذه الاختبارات علامة (*)، وتؤكد المقارنات بين نسب الذكاء كما تتحدد بالمقياس الكلى والمقياس المختصر تطابقا كبيرا بينهما حيث يصل معامل الارتباط بينهما إلى مستوى معامل ثبات المقياس الكلى (٧١، ٧٢).

تطبيق المقياس:

يتطلب هذا المقياس - كغيره من اختبارات الذكاء الفردية - فاحصاً على درجة كبيرة من المهارة والتدريب لأن تطبيق هذا المقياس وتصحيحه على درجة كبيرة من التعقد ويستلزم أن تتوافر فى الفاحص ألفة شديدة به وخبرة طويلة فى التطبيق، لأن تردد الفاحص أثناء التطبيق قد يؤدى إلى عجز عن تكوين علاقة بينه وبين المفحوص، بل وفقدان ثقة المفحوص به، بل أن أدنى تغيير فى صياغة الأسئلة قد يؤدى إلى تغيير مستوى صعوبتها كما تحدد فى عينه التقنين، بالإضافة إلى ما يتطلبه المقياس من تصحيح فورى لاختباراته عقب إعطائها مباشرة، لأن تتابع إعطاء الاختبارات يعتمد على أداء المفحوص فى المستويات السابقة.

ويرى كثير من علماء النفس الإكلينيكيين أن مقياس ستانفورد - بينيه ليس مقياسا مقننا للذكاء فحسب. وإنما هو كذلك نوع من المقابلة الإكلينيكية، لأنه يهيىءفرصة للتفاعل بين الفاحص والمفحوص، ويهيىء مصادر أخرى للدلالات الإكلينيكية لدى الأخصائى الإكلينيكى المتمرس ، بل أن مقياس ستانفورد - بينيه، أكثر من غيره من الاختبارات الفردية، يهيىء الفرصة للفاحص أن يلاحظ الطرق التى يلجأ إليها المفحوص فى العمل، وأساليب حل للمشكلات، وغير ذلك من الأساليب الكيفية للأداء، بل إن الفاحص يمكنه أن يحكم من خلال الموقف الاختبارى الذى يستخدم فيه هذا المقياس على بعض خصائص الشخصية مثل مستوى النشاط والثقة بالنفس والمثابرة والقدرة على التركيز، وبالطبع فإن أى ملاحظات كيفية يمكن تسجيلها أثناء موقف الاختبار يجب تحديدها فى ذاتها، ولايجب تفسيرها على نفس النحو الذى تفسر به درجات الاختبار الموضوعى، وتعتمد قيمة هذه الملاحظات الكيفية إلى حد كبير على مهارة الفاحص وخبرته وإعداده السيكولوجى العميق، بالإضافة إلى وعيه الشديد بمزالق وحدود هذا النوع من الملاحظات المباشرة، وقد استطاع موريارترى (٨٩) أن يسجل أنماط الملاحظات الإكلينيكية فى موقف الاختبار الفردى، على أساس أن جلسة الاختبار هى لدراسة سلوك الطفل أثناء مواجهة تسود فيه عوامل التحدى والاستثارة والصعوبة والإحباط.

وفى أعطاء مقياس ستانفورد - بينيه لا يحاول المفحوص جميع أسئلة المقياس، وإنما يختبر المفحوص فقط حول مدى المستويات العمرية الملائمة لمستواه العقلى، ولايتطلب المقياس أكثر من ٣٠ – ٤٠ دقيقة للأطفال الصغار، ولا يتطلب أكثر من ساعة ونصف للراشدين، والإجراء المقنن فى تطبيق المقياس هو أن نبدأ الاختبار من مستوى أدنى قليلا من العمر العقلى المتوقع للمفحوص وعلى ذلك فإن الاختبارات الأولى يجب أن تكون سهلة بالقدر الذى يستثير فى المفحوص الثقة بالنفس، بشرط ألا تكون سهلة جداً بحيث تؤدى بالمفحوص إلى الاستهانة بالاختبار أو الملل، والإجراء الأكثر موضوعية فى هذا الصدد هو أن نبدأ بأعطاء المفحوص أسئلة حتى يصل إلى العمر الذى يستطيع المفحوص أن يجيب على جميع أسئلته، ويصبح هذا المستوى هو «العمر القاعدى» basal age وبعد ذلك نعطيه أسئلة عمره الزمنى فإن أجاب عليها جميعا أو على بعضها نعطيه أسئلة الأعمار الأعلى منه تدريجيا حتى نصل إلى العمر الذى يفشل المفحوص فى الإجابة على جميع أسئلته ويسمى هذا المستوى «الحد الأعلى للعمر» أو سقف الاختبار ceiling age ، وحين يصل الفاحص إلى هذا المستوى يتوقف عن إعطاء المقياس*.

تصحيح المقياس:

تصحح جميع أسئلة مقياس ستانفورد - بينيه على أساس النجاح أو الفشل (واحد أو صفر)، وتحدد كراسة التعليمات الأداء المطلوب «للنجاح» فى كل سؤال ، وقد تظهر نفس الأسئلة فى مستويات عمرية مختلفة ولكنها تصحح بمستويات مختلفة للنجاح، وفى هذه الحالة تطبق الأسئلة مرة واحدة ثم يحدد أداء المفحوص المستوى العمرى الذى يستحقه، ومن ذلك مثلا أن سؤال المفردات يطبق ابتداء من المستوى العمرى ٦ سنوات حتى المستوى الثالث للراشد المتفوق اعتمادا على عدد الكلمات الذى يستطيع المفحوص تحديدها وتعريفها.

والواقع أن الأسئلة التى ينجح فيها المفحوص أو يفشل تظهر مقدار من لتشتت فى المستويات العمرية المتتابعة، فنحن لا نجد مفحوصاً يجيب على جميع أسئلة عمره

---

* لاحظ الإختلاف فى تحديد العمر القاعدى وسقف المقياس فى طبعة ١٩٨٤ كما تحدد فى وصف هذه الطبعة فى الصفحات السابقة. وايضا فى وصف المقياس ومستوياته فى الشكل الموضح لمحتوى أسئلة.

العقلى أو أدنى منه ويفشل فى جميع الأسئلة الأعلى من هذا المستوى، وإنما نجد فى العادة أن الأسئلة التى ينجح فيها تنتشر فى مستويات عمرية عديدة يحددها العمر القاعدى من ناحية والحد الأعلى للعمر (سقف الاختبار) من ناحية أخرى، ويحسب العمر العقلى فى هذه الحالة بالعمر القاعدى مضافا إليه أى شهور عقلية لأى اختبار اجتازه من مستويات أعلى من عمره القاعدى تبعا للأوزان المبينة فى الجدول رقم(١٨).

جدول رقم (١٨)

يبين أوزان أسئلة مقياس استانفورد ـ بينيه فى الأعمار المختلفة

العمر	عدد الأسئلة	وزن السؤال بالشهر	مجموع التقدير الوزين لكل مستوى بالشهر
٢	٦	١	٦
٢ – ٦	٦	١	٦
٣	٦	١	٦
٣ – ٦	٦	١	٦
٤	٦	١	٦
٤ – ٦	٦	١	٦
٥	٦	١	٦
٦	٦	٢	١٢
٧	٦	٢	١٢
٨	٦	٢	١٢
٩	٦	٢	١٢
١٠	٦	٢	١٢
١١	٦	٢	١٢
١٢	٦	٢	١٢
١٣	٦	٢	١٢
١٤	٦	٢	١٢
الراشد المتوسط	٨	٢	١٦
الراشد المتفوق (١)	٦	٤	٢٤
الراشد المتفوق (٢)	٦	٥	٣٠
الراشد المتفوق (٣)	٦	٦	٣٦
المجموع	١٢٢	المجموع	٢٥٦
	+		+
عدد الأسئلة الاحتياطية	٢٠	العمر الأدنى لبدء الاختبار	١٨
المجموع الكلى	١٤٢	شهرا	٢٧٤
		شهور	سنة
		أى ١٠	٢٢

ومن الوجهة النظرية فإن أعلى عمر عقلى يمكن الوصول إليه بمقياس ستانفورد - بينيه - هو ٢٢ سنة و١٠ شهور، والواقع أن هذه الدرجة لاتدل على عمر عقلى حقيقى، وإنما هى درجة عددية تدل على مقدار الزيادة عن الأداء المتوسط الذى يصدر عن الراشد العادى، وهى بالتأكيد لا تطابق ما يصل إليه الشخص المتوسط من عمر زمنى مقداره ٢٢ سنة و١٠ شهور، لأن الشخص العادى يحصل على عمر عقلى مقداره ١٦ سنة و٨ شهور كما تحدده معايير ١٩٧٢(*)، أما بالنسبة لأى راشد يزيد عمره عن ١٨ سنة فإن أى عمر عقلى مقداره ١٦ سنة، ٨ شهور يحصل على نسبة الذكاء ١٠٠.

ويجب أن نؤكد هنا أنه بالنسبة للمفحوصين ابتداء من مرحلة المراهقة (١٢ +) لا يعطى العمر العقلى فى هذا المقياس نفس الدلالة التى يعطيها فى الأعمار الأدنى، لأنه بعد سن ١٣ يقل متوسط الأعمار العقلية عن متوسط الأعمار الزمنية، ولذلك فإن مقياس ستانفورد - بينيه لايلائم الراشدين وبخاصة العاديين والمتفوقين منهم. ورغم وجود ٣ مستويات للراشد المتفوق فيه فإن «سقف» الاختبار لايلائم معظم الراشدين المتفوقين أو حتى الراشدين العاديين كما تؤكد دراسات كيندى وزملائه (٨٦) فمن المستحيل أن نجد فى المقياس سقفا مناسبا لهؤلاء، أى مستوى عمرى يعجز المفحوصون فى الإجابة عن جميع أسئلته وبالإضافة إلى ذلك فإن مقياس ستانفورد - بينيه أكثر جاذبية للصغار منه للمراهقين والراشدين من حيث محتواه.

ومن أهم الإضافات إلى طبعة عام ١٩٦٠ إحلال نسبة الذكاء الانحرافية محل نسبة الذكاء العادية التى استخدمت فى المقياس منذ ظهوره عام ١٩١٩. وهذه النسبة الجديدة عبارة عن درجة معيارية معدلة متوسطها ١٠٠ وانحرافها المعيارى ١٦. والميزة الأساسية فى هذا النوع من المعايير أنها تعطى درجات قابلة للمقارنة فى مختلف الأعمار، وبهذا أمكن التغلب على مشكلة عدم استقرار نسبة الذكاء العادية، فعلى الرغم من الجهود التى بذلت عام ١٩٣٧ للحصول على تشتتات ثابتة لنسبة الذكاء فى مختلف الأعمار، فإننا نلاحظ أن هذا المقياس تذبذبت فيه الانحرافات المعيارية لمختلف الأعمار من انحراف معيارى ١٣ (فى سن ٦) إلى انحراف معيارى ٢١ (فى سن ٢ - ٦) ومعنى هذا أن نسبة الذكاء ١١٣ فى سن ٦ تقابل نسبة الذكاء

(*) كان العمر ١٦ سنة فى طبعة ١٩٣٧.

١٢١ فى سن ٢ – ٦، وقد وضع مكنمار جداول خاصة بتصحيح آثار هذه الاختلافات فى طبعة ١٩٣٧ (فى ٧٨) وتتلخص فكرة التصحيح فى أنه إذا كان العمر العقلى للمفحوص بين سن ١٣، ١٦ فإنه يجب تعديل العمر الزمنى الذى يؤخذ فى الاعتبار عند حساب نسبة الذكاء، وذلك بأن نضيف ثلثى الشهور الإضافية التى عاشها المفحوص بعد سن ١٣، فمثلا إذا كان العمر الزمنى الحقيقى ١٤ سنة فإنه يحسب ١٣ – ٨، وإذا كان ١٥ سنة يحسب ١٤ – ٤.

وقد أمكن التغلب على كل هذه الصعوبات فى طبعة ١٩٦٠ باستخدام نسبة الذكاء الانحرافية والتى تعنى استخدام انحراف معيارى موحد لجميع الأعمار، وقد أعد بينو (٧٢) جداول تستخرج منه مباشرة نسبة الذكاء الانحرافية من كل من العمر العقلى والعمر الزمنى بالسنوات والشهور، وقد نشرت هذه الجداول فى كراسة تعليمات المقياس الصادرة عام ١٩٧٣ (١١١).

وتعتمد فكرة التحويل فى جداول بينو على اعتبار نسبة الذكاء التقليدية هى الدرجة الخام للمفحوص تحول إلى نسبة ذكاء انحرافية باستخدام متوسطات وانحرافات معيارية تجريبية لنسب الذكاء التقليدية لكل فئة من الفئات العمرية، ثم يستخدم الانحراف المعيارى الجديد ١٦ والمتوسط الجديد ١٠٠، أى معادلة نسبة الذكاء الانحرافية فى مقياس ستانفورد ـ بينيه هى :

نسبة الذكاء الانحرافية =

$$\left[\left(\frac{\text{نسبة الذكاء التقليدية ـ المتوسط التجريبى}}{\text{الانحراف المعيارى التجريبى}}\right) \times ١٦\right] + ١٠٠$$

أو

$$\left[(\text{نسبة الذكاء التقليدية ـ المتوسط التجريبى}) \times \frac{١٦}{\text{الانحراف المعيارى}}\right] + ١٠٠$$

وقام بينو بقسمة ١٦ على الانحراف المعيارى التجريبى لكل فئة عمرية فحصل على ما أسماه بالمقدار الثابت ك K وأعد جدولا بهذه الثوابت لفئات عمرية مداها شهر ، والجدول رقم (١٩) يوضح هذه الثوابت لفئات العمر المختلفة إلا أن ما يجب أن نشير إليه هو أن حساب العمر العقلى لايزال يتم بالطريقة المستخدمة منذ عام ١٩٣٧ ، أى أن التعديلات أدخلت على مفهوم نسبة الذكاء التقليدية دون أن تمس مفهوم العمر العقلى ذاته .

جدول (١٩) يبين الثوابت اللازمة لتحويل نسبة الذكاء التقليدية إلى نسبة ذكاء انحرافية .

العمر الزمنى	المتوسط التجريبى	ك	العمر الزمنى	المتوسط التجريبى	ك
٢	١٠٢	,٩٩			
٢ – ٦	١٠٣	,٨٧	٨	١٠٢	١,٠٢
٣	١٠٤	,٨٨	٩	١٠٣	,٩٨
٣ – ٦	١٠٢	,٩٢	١٠	١٠٣	,٩٤
٤	١٠٢	,٩٨	١١	١٠٢	,٩١
٤ – ٦	١٠١	١,٠٤	١٢	١٠٢	,٨٨
٥	١٠٠	١,٠٩	١٣	١٠١	,٨٩
٦	١٠٠	١,١١	١٤	١٠١	,٩٠
٧	١٠١	١,٠٧	الراشد	١٠٥	,٩٠

ويوجد تعديل آخر أضافته طبعة ١٩٦٠ وأكدته طبعة ١٩٧٣ ، واعتمد على النتائج التى تؤكد أن التحسن فى الأداء فى المقياس يستمر إلى سن ١٨ سنة بدلا من التوقف عند سن ١٦ كما كان الحال فى طبعة ١٩٣٧ ، فقد بينت البحوث الطولية فى النمو العقلى التى أجريت خلال الربع قرن بين طبعتى المقياس أن القدرات العقلية كما تقاس باختبارات الذكاء تستمر فى التحسن والارتقاء أبعد مما كان مفترضا (٢٨) ، وأكثر الأدلة مباشرة على زيادة التحسن فى مقياس ستانفورد - بينيه بعد سن ١٦ قدمته بحوث برادوى وطومسون وكرافنر (٨٢) ، بعد انقضاء ١٠ سنوات ثم بعد انقضاء ٢٥

سنة، فلوحظ أن هؤلاء المفحوصين حين أختبروا أول الأمر وكانت أعمارهم الزمنية ٢ – ٥,٥ سنة ثم اختبروا بعد ١٠ سنوات، أن متوسطهم فى نسبة الذكاء لم يختلف فى الحالتين عما حددته معايير المقياس ولكن هذا المتوسط أظهر ارتفاعا له دلالة إحصائية بمقدار ١١,٣ نقطة بين الاختبار الثانى والثالث (أى بعد انقضاء ١٥ سنة من الاختبار الثانى و٢٥ سنة من الاختبار الأول). وهذه الزيادة نتجت عن أن المفحوصين استمروا فى التحسن بعد سن ١٦ سنة، بينما نجد أن حساب نسب الذكاء الذى اعتمدت عليه طبعة ١٩٣٧ افترضت توقف النمو العقلى عند هذا السن.

وعند المقارنة بين طبعات ١٩٣٧، ١٩٧٢، ١٩٨٥ نجد أن المعايير فى الطبعات الأحدث تعتمد على عينات أكثر تمثيلا، كما أنها أكثر حداثة وتعكس أى تغيرات فى الظروف الحضارية والثقافية يمكن أن تؤثر فى الأداء الاختبارى، ومن الطريف أن نذكر أن المعايير الأحدث تظهر بعض التحسن فى أداء الاختبار فى جميع الأعمار، ويكون هذا التحسن أكثر جوهرية فى أعمار ما قبل المدرسة يصل فى متوسطه إلى حوالى عشر نقاط، ويرجع مؤلفو الاختبار هذا التحسن إلى أثر وسائل الإعلام من ناحية وإلى ارتفاع المستوى الثقافى والتعليمى للوالدين[1].

طريقة حساب العمر العقلى ونسبة الذكاء التقليدية:

يمكن توضيح الأسلوب الذى يستخدم فى الحصول على العمر العقلى ونسبة الذكاء التقليدية من مقياس ستانفورد ـ بينيه بالمثال الموضح فى الجدول رقم (٢٠).

(١) راجع العرض لطبعة ١٩٨٤ .

جدول رقم (٢٠)

يوضح طريقة حساب العمر العقلى ونسبة الذكاء التقليدية

العمر	عدد الأسئلة التى أجاب عليها المفحوص	وزن السؤال بالشهر	المجموع الكلى للأوزان	
		طفل عمره الزمنى ٦ سنوات، ٤ شهور	شهر	سنة
٤	٦	العمر القاعدى	٦	٤
٤ – ٦	٥	١	٥	–
٥	٣	١	٣	–
٦	٣	٢	٦	–
٧	٢	٢	٤	–
٨	١	٢	٢	–
٩	صفر	سقف الاختبار	–	–
			٢٠	٤

العمر العقلى = ٨ شهر ٥ سنة

العمر الزمنى = ٤ شهر ٦ سنة

نسبة الذكاء التقليدية = $\frac{\text{العمر العقلى}}{\text{العمر الزمنى}} \times ١٠٠ = \frac{\text{٦٨ شهرا}}{\text{٧٦ شهرا}} \times ١٠٠ = ٨٩$

وبالطبع تحول نسبة الذكاء التقليدية ٨٩ إلى نسبة ذكاء انحرافية باستخدام جداول بينو المشار إليها.

تفسير معايير المقياس:

من أهم مزايا مقياس ستانفورد - بينيه ما يتوافر حوله من بيانات امبريقية وخبرة كلينيكية تراكمت طوال ثمانين عاما منذ مقياس بينيه المبكر. وقد أصبح هذا المقياس عند كثير من علماء النفس الكليكينيين والتربويين وغيرهم من المهتمين بالتقويم النفسى مرادفا للذكاء، وأصبح تحديد مستويات الذكاء يعتمد فى جوهره على نسب الذكاء كما يحددها المقياس ، رغم ما فى هذا الرأى من غلو لا يتسع المقام لمناقشته، والجدول رقم (٢١) يوضح التوزيع والتصنيف النظريين لنسب الذكاء الانحرافية لطبعتى ١٩٦٠، ١٩٧٣.

جدول رقم (٢١)

يبين التوزيع والتصنيف النظرى لنسب الذكاء الانحرافية لمقياس ستانفورد ـ بينيه

الدرجة المعيارية	نسب الذكاء الانحرافية	التصنيف	النسب المئوية للحالات
٣ –	٥٢ ٦٤	التخلف العقلى (الضعف العقلى)	١,٢ ٪
٢ –	٦٨ ٧٦ ٨٤	الحالات الهامشية (قد تصنف أغبياء أو ضعاف عقول	٥,٥ ٪
١ –	٨٨ ١٠٠	الغباء (بطء التعلم)	١٦ ٪
صفر	١١٢ ١١٦	المتوسط (أو العادى)	٥٤,٧ ٪
١ +	١٢٤ ١٣٢ ١٤٨	المتفوقون	١٦ ٪
٢ + ٣+		المتفوقون جداً	٦,٥ ٪
أعلى من ذلك		العبقرية أو شبه العبقرية	٠,١ ٪

وقد كان من المعتاد لعدة سنوات استخدام نسبة الذكاء التقليدية (٧٠) باعتبارها نقطة قطع تقريبية للتخلف العقلى أو الضعف العقلى المحتمل، وتتطابق هذه الفئة مع النسبة ١ ٪ أو ٢ ٪ من الحد الأدنى للتوزيع، أما الجدول فيوضح لنا أن نقطة القطع هذه أصبحت عند الدرجة المعيارية (–٢,٢٨) والتى تتطابق مع نسبة الذكاء الانحرافية ٦٤ والتى يقع أدنى منها ١,٢ ٪ من الحالات.

وعموما فإن الفضل يعود إلى هذا المقياس فى استخدام التقسيم الثلاثى الشهير للضعف العقلى: المأفونون والبلهاء والمعتوهون، وهى ألفاظ لم تستخدم فى الوقت الحاضر وحل محلها ما يلى:

١ – الحالات القابلة للتعلم educable وتمتد نسب ذكائها بين ٥٠، ٧٥ ويمكنها أن تتعلم حتى الصف الثالث فى المرحلة الابتدائية ولايتجاوز حدها الأقصى فى التعلم نهاية المرحلة الابتدائية إذا تعلمت موقف تعليمى خاص.

٢ – الحالات القابلة للتدريب trainable وتمتد نسب ذكائها بين ٢٥، ٥٠ ويمكنها أن تتعلم رعاية الذات والتوافق الاجتماعى فى بيئة حماية خاصة.

٣ – الحالات التى تحتاج إلى الإيداع فى المؤسسات custodial وتقل نسب ذكائها عن ٢٥ وتتطلب رعاية وحماية خاصة.

والتصنيف الراهن للضعف العقلى يقسم هذه الحالات إلى أربع فئات يوضعها الجدول رقم (٢٢).

جدول رقم (٢٢)

يبين مستويات الضعف العقلى فى صورة نسب ذكاء انحرافية

نقاط القطع (وحدات ع)	نقاط القطع (وحدات ع)	مدى نسب الذكاء والانحراف (ع = ١٦)	النسب المئوية للحالات
الضعف العقلى الخفيف	– ٢	٥٢ – ٦٨	٢,١٤
الضعف العقلى المتوسط	– ٣	٣٦ – ٥١	,١٣
الضعف العقلى الشديد	– ٤	٢٠ – ٣٥	,٠٠٣
الضعف العقلى الخطير	– ٥	١٩ فأدنى	,٠٠٠٠٣

والميزة الجوهرية فى هذا التصنيف أنه يصلح للتحويل إلى درجات معيارية أو نسب ذكاء انحرافية فى أى مقياس للذكاء.

والواقع أن شيوع تصنيف مستويات نسب الذكاء ـ رغم فائدته فى تقنين تفسير الأداء الاختبارى ـ يتضمن بعض المخاطر، فهو كغيره من صور التصنيف السلوكى يجب أن يستخدم بمرونة كافية، كما يجب ألا يستبعد البيانات الأخرى التى تتوافر عن المفحوص، وبالطبع لا يوجد حد فاصل تماما بين فئات التصنيف، ومن ذلك مثلا لا يوجد خط يميز تماما بين التخلف العقلى والحالات الهامشية، أو بين التفوق والتفوق الشديد، فقد لوحظ مثلا أن بعض المفحوصين من ذوى نسب الذكاء ٦٠ يتكيفون تكيفا عاديا لمطالب الحياة اليومية، بينما نجد بعض الأشخاص من ذوى نسب الذكاء ١٠٠ يتطلبون الإيداع فى المؤسسات ، كما أن بعض الأشخاص من ذوى نسب الذكاء ١٦٠

قد لا يعيشون حياة متميزة، بينما نجد البعض الآخر من ذوى نسب الذكاء الأقرب إلى ١٠٠ يسهمون بعض الإسهامات الممتازة، ومعنى ذلك أن القرارات التى تتحدد حول تشخيص الضعف العقلى أو التفوق العقلى يجب ألا تعتمد على نسب الذكاء وحدها، وإنما يجب فى حالة الضعف العقلى والقرارات المتصلة به كالإيداع فى المؤسسات أو الالتحاق بمدارس التربية الفكرية الاهتمام بمتغيرات أخرى مثل النضج الاجتماعى والتوافق الانفعالى والظروف الصحية والجسمية والأحوال الأسرية والظروف المدرسية وغير ذلك من العوامل التى تؤثر فى الأداء العقلى وفى جميع الحالات يجب الاهتمام بكل من النشاط العقلى للفرد ونضجه الاجتماعى وسلوكه التوافقى جميعا وفى وقت واحد، وبالمثل فإن نسب الذكاء العالية لا ترادف العبقرية، لأن الإنجاز الرفيع قد يتطلب بالإضافة إلى ذلك الابتكار والمواهب الخاصة والمثابرة والاهتمام بالهدف وغيرها من العوامل العقلية والانفعالية والدافعية (٣٩).

ثبات المقياس:

أجريت دراسات الثبات على طبعة ١٩٣٧ من المقياس باستخدام طريقة الصور المتكافئة على فترات مقدارها أسبوع أو أقل (للتحقق من مصدرى تباين الخطأ: الاستقرار الوقتى وتكافؤ المحتوى)، وقد لوحظ أن المقياس يميل إلى أن يكون أكثر ثباتا فى الأعمار الكبيرة منه فى الأعمار الصغيرة، وفى مستويات نسب الذكاء الدنيا عنه فى نسب الذكاء العليا.

والواقع أن زيادة الثبات مع الزيادة فى العمر هى من خصائص الاختبارات النفسية بوجه عام، وتنتج جزئيا عن التحكم الأفضل فى ظروف إعطاء الاختبار فى حالة المفحوصين كبار السن (وبخاصة حين نقارنهم بأطفال ما قبل المدرسة)، بالإضافة إلى العوامل الأخرى المسهمة فى زيادة معاملات الثبات ومنها بطء معدل النمو مع التقدم فى العمر.

أما عن ارتفاع معاملات الثبات فى حالة نسب الذكاء الدنيا فى أى عمر زمنى فترتبط فيما يبدو بخصائص مقياس ستانفورد - بينيه على وجه الخصوص، فقد أشرنا إلى أن الأوزان النسبية لأسئلة الأعمار المختلفة تختلف، فهى فى الأعمار الدنيا مقدارها شهر، وفى الأعمار المتوسطة شهران، وفى المستويات العليا قد تكون ٤ شهور أو ٥ شهور أو ٦

شهور، ومعنى ذلك أن طريقة وزن الدرجات تميل إلى «تضخيم» أخطاء القياس فى المستويات العليا، لأن النجاح أو الفشل الناجم عن الصدفة فى سؤال واحد يؤدى إلى فرق أكبر فى الدرجة الكلية فى هذه المستويات عنه فى المستويات الدنيا.، وحيث أن المفحوصين من ذوى نسب الذكاء العالية (من أى عمر زمنى) عادة ما يجيبون على الأسئلة التى تقع فى المستويات العليا من المقياس فإن نسب ذكائهم نتيجة لذلك تتضمن مقدارا أكبر من أخطاء القياس وبالتالى تنخفض معاملات الثبات.

وتتأكد هذه الحقيقة من ملاحظة جدول انتشار درجات الصورتين م، ل من طبعة ١٩٣٧ الذى تأكد أنه شكل المروحة فى أى عمر وهذا يعنى أن الصورتين ترتبطان عاليا فى المستويات الدنيا من نسب الذكاء، أما فى المستويات العليا فإن العلاقة بين الصورتين تتناقص بوضوح.

وفى مثل هذه الأحوال يكون من المفضل حساب معامل ارتباط كلى (كمؤشر للثبات)، ولهذا حسبت معاملات ثبات مختلفة للفئات المختلفة من نسب الذكاء فى كل مرحلة عمرية.

وبوجه عام يمكن القول أن مقياس ستانفورد ـ بينيه على درجة عالية من الثبات فى مختلف المستويات العمرية وفى مختلف نسب الذكاء، بالرغم من تجانس مجموعات عينة التقنين من حيث العمر الزمنى، مما يؤثر بالطبع فى معاملات الارتباط عموما، ويؤدى إلى نقصانها.. وقد تراوحت هذه المعاملات بالنسبة للأعمار ٢,٥ – ٥,٥ بين ٠,٨٣ لذوى نسب الذكاء العالية، ٠,٩١ لذوى نسب الذكاء المنخفضة، وبالنسبة للأعمار من ٦ – ١٣ تتراوح هذه المعاملات بين ٠,٩١ ، ٠,٩٧ ولنفس المستويات من نسب الذكاء. كذلك تم استخدام طريقة اعادة التطبيق على عينة قوامها ٥٧ طفلا أعمارهم خمس سنوات، وخمس وخمسون طفلا أعمارهم ٨ سنوات وقد تم اعادة التطبيق لمدة تتراوح بين شهرين وثمانية أشهر وجاءت معاملات الثبات ٩١, ، ٩٠, للدرجة الكلية للمقياس للمجموعتين على التوالى. أما بالنسبة لمعاملات الثبات للاختبارات الفرعية فقد كان هناك تذبذبا لقيمها. ويتضمن دليل الاختبار بيانات وافية عن معاملات الثبات فى العينات المختلفة، وبالنسبة للأعمار من ١٤ – ١٨ كانت معاملات الثبات المطابقة للمستويين هى ٠,٩٥ ، ٠,٩٨.

أما عن تحديد معاملات الثبات لطبعة ١٩٨٤ فقد تم بحساب معاملات الإتساق الداخلى وإعادة التطبيق، وكذلك استخدام معادلة كيودر – ريتشاردسون وذلك على عينات قوامها ٥٠٠٠ (خمسة الاف مفحوصا) يتراوح عمرها الزمنى بين عامان وثلاثة وعشرون عاما تم اختيارها من ٤٧ ولاية أمريكية، وقد تم مراعاة المناطق الجغرافية والخصائص الديموجرافية بنسبتها التى ظهرت فى تعداد ١٩٨٠ للولايات المتحدة الأمريكية. وقد تراوحت معامات الثبات بين ٩٥, و ٩٩, للدرجات الكلية كذلك كانت معاملات الارتباط مرتفعة بين المجالات المعرفية الأربعة حيث تراوحت معاملات الثبات بين ٨٠, – ٩٠, فيما عدا اختبار الذاكرة فقط تراوحت بين ٦٦, – ٧٨,، وبصفة عامة كانت معاملات الثبات أعلى فى الاعمار الأكبر.

صدق المقياس:

(أ) الصدق المرتبط بالمحكات: تتوفر بيانات كافية عن الصدق المرتبط بالمحكات سواء كان صدقا تلازميا أو صدقا تنبؤيا فى ضوء محك التحصيل المدرسى، فمنذ مقياس ستانفورد- بينيه عام ١٩١٦ لم يتوقف الباحثون عن حساب معاملات الارتباطات بينه وبين الدرجات المدرسية وتقديرات المعلمين ودرجات الاختبارات التحصيلية المقننة، ويتراوح معظم هذه المعاملات بين ٠,٤٠ ، ٠,٧٥ كما ارتبط بالتفوق والتخلف الدارسيين بوجه عام.

وقد وجد أن هذا المقياس- كغيره من مقاييس الذكاء- يرتبط ارتباطا عاليا بالأداء فى جميع المواد الدراسية الأكاديمية تقريبا، إلا أن ارتباطه أعلى ما يكون بالمقررات اللغوية (كمواد اللغات أوالمواد الاجتماعية) . وقد تأكدت هذه النتائج عند اختبار التلاميذ والطلاب فى مختلف المراحل التعليمية وبالطبع فإن معاملات الارتباط تميل إلى التناقص مع ارتفاع السلم التعليمى بسبب عوامل الانتقاء وعدم ملاءمة الاختبار مما يؤدى إلى نقصان التباين بين المفحوصين.

وفى تفسير نسبة الذكاء يجب أن نؤكد أن مقياس ستانفورد- بينيه (كغيرة من اختبارات الذكاء العام) هو فى الأغلب مقياس للاستعداد المدرسى أو التحصيل الأكاديمى العام، ومشبع إلى حد كبير بالمحتوى اللفظى، وبخاصة فى «اختبارات المستويات العمرية العليا، ولذلك فإن الأفراد من ذوى النقائص اللغوية، أو أولئك الذين

يتفوقون فى بعض القدرات غير اللغوية (كالقدرات الموسيقية) يحصلون على درجات منخفضة نسبياً فى المقياس، وكذلك فإنه توجد بعض المجالات لا يلعب فيها الاستعداد المدرسى أو الفهم اللغوى دوراً مهمًا. وفيها لا تصلح اختبارات الذكاء من نمط ستانفورد - بينيه، ومعنى ذلك أن استخدام الاختبارات بصفة عامة يجب أن يكون فى المواقف التى تلائمها، وإلا فإن الباحث يتوقع من المقياس ما لا يعطيه.

(ب) صدق التكوين الفرضى: من أهم المحكات التى استخدمت فى انتقاء اختبارات مقياس ستانفورد - بينيه محك تمايز العمر، ومعنى ذلك أن المقياس يقيس قدرات متزايدة مع العمر منذ الطفولة حتى الرشد (فى ظل إطار ثقافى وحضارى معين بالطبع).

وقد استخدم أيضًا محك الاتساق الداخلى فى انتقاء الأسئلة. وقد أكدت النتائج وجود نوع من التجانس الوظيفى فى مقياس استانفورد - بينيه، بالرغم من التنوع الظاهرى فى المحتوى، والدليل على هذا أن متوسط معاملات الارتباط بين السؤال والمقياس ككل هو ٠,٦٦ فى طبعة عام ١٩٦٠، إلا أن من الواضح غلبة الطابع اللفظى على المقياس، ويتضح هذا من ارتفاع معامل الارتباط بين الأسئلة اللفظية والأداء الكلى فى المقياس عنه فى حالة الأسئلة غير اللفظية.

ومن البراهين الإضافية على صدق التكوين الفرضى نتائج الدراسات العاملية العديدة التى أجريت على اختبارات مقياس ستانفورد - بينيه، ويعتمد هذا البرهان على أنه لو كانت نسب الذكاء قابلة للمقارنة فى مختلف الأعمار، فإن المقياس يجب أن يكون له نفس التكوين العاملى فى جميع مستويات الأعمار، بل يجب أن يكون المقياس مشبعا بعامل عام واحد حتى يمكن تفسير نسبة الذكاء تفسيرا واضحًا، أما إذا كان المقياس مشبعا بعاملين طائفيين أو أكثر فإن نسبة الذكاء الواحدة التى يحصل عليها أفراد متعددون (نسبة ذكاء ١٢٠ مثل) قد لا تدل على معنى سيكولوجى واحد.

وقد تحقق هذا الغرض من نتائج التحليلات العاملية التى أجراها ماكنمر (٨٨) لاختبارات مقياس ستانفورد - بينيه فى ١٤ مستوى عمريا ابتداء من سن سنتين وحتى سن ١٨ سنة، وكان عدد المفحوصين فى كل تحليل عاملى يتراوح بين ٩٩، ٢٠٠، كما تراوح عدد الأسئلة التى خضعت للتحليل بين ١٩، ٣٥، وفى كل الأحوال حسبت

معاملات الارتباط الرباعية بين المفردات، وخضعت مصفوفات الارتباط للتحليل العاملى. وتؤكد نتائج ماكنمر، أن الأداء فى مقياس ستانفورد – بينيه يمكن تفسيره فى ضوء عامل عام واحد كما وجدت عوامل طائفية إضافية فى بعض المستويات العمرية، إلا أن اسهام هذه العوامل فى التباين كان قليلا، كما ثبت أن العامل العام المشترك الذى وجد فى المستويات العمرية المتقاربة كان متشابها إلى حد كبير، ومن أهم النتائج التى توصل إليها ماكنمر أن العامل العام يصبح أكثر لفظية كلما تقدمت الأعمار، ومن ذلك مثلا أن تشبع العامل العام باختبار المفردات ارتفع من ٠,٥٩ فى سن ٦ سنوات إلى ٠,٩١ فى سن ١٨ سنة.

وأجرى جونز (٨٤، ٨٥) دراستين عامليتين لأسئلة مقياس ستانفورد - بينيه على ٤ مجموعات من الأطفال أعمارهم ٧، ٩، ١١، ١٣ سنة على التوالى، وكانت كل مجموعة تتكون من ١٠٠ ذكر، ١٠٠ أنثى، ووجد أنه فى كل عمر يوجد عدد من «القدرات» المتمايزة ولكنها مرتبطة، ومنها العوامل اللفظية والتذكرية الاستدلالية والمكانية والإدراكية، وقد اختلفت هذه العوامل وأوزانها النسبية من مستوى عمرى لآخر، وتتفق هذه النتيجة مع نتائج دراسات عاملية أخرى على هذا المقياس يمكن الرجوع إليها ملخصة فى (٢٤).

والواقع أن الأساليب الإحصائية التى استخدمها ماكنمر وجونز من النوع الذى يؤكد دور العامل العام فى الحالة الأولى (دراسات ماكنمر)، ودور العوامل الطائفية المختلفة فى الحالة الثانية(دراسات جونز)، وحين نتأمل النتائج التى توصل إليها الباحثان نجد أنها تؤكد وجود قدر مشترك فى القياس ككل، وهى خاصية من خصائص بناء المقياس حسب طريقة انتقاء أسئلته التى كانت تؤكد ارتفاع معاملات الارتباط بين السؤال والدرجة الكلية فى المقياس وبالإضافة إلى ذلك فإن الأداء يتأثر كذلك بعدد من القدرات المتخصصة التى يختلف تكوينها تبعا لمستوى العمر الذى نختبره.

وقد أجريت عدة دراسات لتحديد صدق طبعة ١٩٨٤ واستخدمت فى ذلك طرقا متعددة ومعالجات مختلفة حيث تم الإستفادة من الدراسات التى أجريت على الطبعات المختلفة للمقياس. وقد تم البدء فى تحديد الصدق من خلال دراسة بينية للمفردات فى

المقياس لتتلاءم والبنية الكلية للمقياس، كما تم دراسة النتائج التى تم التوصل لها من تطبيق المفردات وتحليل ما تم التوصل له من نتائج من الدراسات الميدانية لتجريب الأنواع المختلفة من المفردات . وقد تم تحديد الصدق بالطرق الآتية:

أ- التحليل العاملى

ب- معاملات الارتباط مع اختبارات أخرى للذكاء.

جـ- تطبيق الإختبارات على عينات ذات خصائص محددة مسبقا (الموهوبون / ضعاف العقول) (ثورنديك وآخرون ١١٤ : ١٩٨٦) .

وعند استخدام التحليل العاملى لتحديد الصدق تم التوصل الى عامل عام وثلاثة عوامل طائفية لمجموعات الاختبارات فيما عدا مجموعة اختبار الاستدلال البصرى. كما وجدت عوامل خاصة للإختبارات الفرعية. (ويتضمن دليل الإختبار تفصيلا لكل هذه العوامل) .

وباستخدام طريقة حساب الصدق من خلال حساب معاملات الإرتباط مع اختبارات الذكاء الأخرى فقد تم الحصول على معاملات الارتباط بين الأداء فى الطبعة الثالثة والطبعة الرابعة للمقياس وبلغت قيمة معامل الارتباط بين الدرجة الكلية من المقياس الرابع (Composite Score) ونسبة الذكاء للصورة (ل ، م) ٨١, .

وباستخدام الطريقة الثالثة (محك المجموعات المتضادة) تم الحصول على فروق دالة بين المجموعات لصالح المجموعة الفائقة، وكانت متوسطات المجموعات الأخرى أقل من متوسط مجموعات التقنين، بينما كان متوسط درجات المجموعة الفائقة أعلى من متوسطات عينات التقنين. (آناستازى واوربانا ١٩٩٧ : ٢١٥) .

أما عن الدراسات التى أجريت فى مصر لتحديد المعالم السيكومترية لمقياس ستانفورد- بينيه فإنها قليلة، وتتلخص فى دراسة قام بها لويس كامل مليكة عام ١٩٦٠ (٤١) للتحقق من صدق الترتيب المبدئى لاختبارات المفردات فى المقياس، وذلك بتطبيقها على مجموعات كبيرة من الأفراد ومن مختلف فئات السن (بين ٣ سنوات، ٣٠ سنة) بلغ عددها ١٦٤٢ مفحوصاً، وفى عام ١٩٦٣ قام محمود الهوارى (٣٠) بدراسة عن ثبات المقياس وصدقه، استخدم فى حساب الثبات طريقة إعادة الاختبار

على ٩٥ تلميذا بالسنة الثالثة الإعدادية (متوسط أعمارهم ١٤ عاما)، بفاصل زمنى بين شهر وشهرين، وبلغ معامل الارتباط ٠,٨٣ أما عينة الصدق فتألفت من ١٦٠ تلميذا بالسنة الثالثة الإعدادية أيضاً، وقد حسب لهم معاملات الارتباط بين درجات المقياس (نسبة الذكاء) وحالات الرسوب والنجاح فى امتحان الشهادة الإعدادية من ناحية، والمجموع الكلى لدرجات هذا الامتحان من ناحية ثانية، ودرجات مواد هذا الامتحان من ناحية ثالثة، وقد وجد هذا الباحث أن معامل الارتباط الثنائى بين نسب الذكاء وحالات النجاح والرسوب هو ٠,٥٤، ومعامل الارتباط بين نسب الذكاء والمجموع الكلى لامتحان الشهادة الإعدادية ٠,٨١٨ أما معاملات ارتباط المقياس بدرجات المواد الدراسية فكانت فى المتوسط ٠,٥٠ ولم يظهر التحصيل اللغوى (فى اللغة العربية) معاملات ارتباط أعلى من التحصيل اللغوى (فى الرياضيات) كما هو الحال فى الدراسات السابقة ومنها بحث بوند (٦٤) الشهير.

وأعد الباحث بعد ذلك مصفوفة معاملات ارتباط درجات الذكاء ودرجات المواد الدراسية المختلفة وأخضعها للتحليل العاملى فوجد أن تشبع مقياس ستانفورد - بينيه بالعامل العام ٠,٧٨٤ وهو أعلى التشبعات بهذا العامل، إلا أن هذا ليس برهانا على صدق التكوين الفرضى للمقياس لأن باقى المتغيرات كانت فى الواقع درجات للتحصيل المدرسى، ولم تتضمن المصفوفة أى مقياس آخر للذكاء.

وقد ظهرت على نسق الصورة المصرية للمقياس (طبعة ١٩٣٧) صورتان عربيتان أخريان إحداهما أردنية أعدها عبدالله زيد الكيلانى (٢٦) والأخرى سعودية أعدتها آمال أحمد مختار صادق (٩)، كما ظهرت صورة كويتية لطبعة ١٩٦٠ أعدها رجاء محمود علام وكمال إبراهيم مرسى (٢٢) كما قام عبدالله عبد العزيز المعيلى (٢٧) بتقنين طبعة ١٩٦٠ على عينة سعودية من ١٧٥ تلميذا بالمرحلة الابتدائية بالرياض تمتد أعمارهم بين ٦ و ١٢ سنة، وحسب له الصدق بمحك اختبار رسم الرجل والذى سبق تقنينه على البيئة السعودية وتراوحت معاملات الارتباط بين ٠,٥٧، ٠,٨٤ للأعمار المختلفة، أما الثبات فحسب بطريقة إعادة الاختبار - بفاصل زمنى طوله أسبوعان، وتراوح معامل الارتباط بين ٠,٨٨ ، ٠,٩١ وحسبت له معايير فى صورة متوسطات عمرية تدل على العمر العقلى، والسؤال الذى لم نجد له إجابة فى هذا البحث هو: كيف استخدم الباحث متوسطات الأعمار للدلالة على العمر العقلى

فى مقياس ستانفورد ـ بينيه، ومن المعلوم أن حساب العمر العقلى فى هذا المقياس يقوم على أسس عرضناها بالتفصيل فيما سبق لا تعتمد على فكرة حساب المتوسط والانحراف المعيارى، ومن الطريف أن الباحث استخدم ضمن معاييره نسبة الذكاء بالمعنى التقليدى، (٢٧ : ١٢٩) فى الوقت الذى انقرض فيه هذا المعيار تماما من التراث السيكومترى المعاصر.

وتكاد تكون الدراسة العربية الوحيدة ـ فى حدود علمنا ـ التى قامت بتقنين طبعة ١٩٧٢ من هذا المقياس هى دراسة جمال محمد على (٢٠)، وبدأ بالصورة المصرية القائمة والتى أعدها محمد عبدالسلام أحمد ولويس كامل مليكة فطبقها على عينة من تلاميذ مرحلة التعليم الأساسى فى مصر (ن = ٢٠٧) تتراوح أعمارهم بين ٦، ١٥ سنة، ثم طبق على الأسئلة الطرق الإحصائية لتحليل الاحتمالات الاختيارية لكل سؤال، وتحليل قائمة الكلمات وترتب على ذلك تغيير وتعديل وحذف وإضافة بعض المفردات لكى تناسب طبيعة الثقافة والبيئة المصرية، واستعان فى كل ذلك بطبعة ١٩٧٢ من المقياس، وانتهى به الأمر إلى إعداد صورة مقترحة للمقياس طبقها على عينة أخرى من نفس المرحلة (ن = ١١٦) للتأكد من صلاحيتها للاستخدام وتلا ذلك تطبيقها على عينة ثالثة (ن = ١٠٠) لحساب الخصائص السيكومترية لها، فحسب الثبات بطريقة كيودر ـ ريتشاردسون فبلغ معامل الثبات ٠,٧٤ كما قدر الصدق بحساب معاملات الارتباط بين كل من الدرجات الخام ونسب الذكاء الانحرافية لها فتراوحت معاملات الارتباط بين ٠,٢٠، ٠,٥٥ وهى معاملات دالة إحصائيا بالرغم من انخفاضها الظاهر، وقد قدر الصدق أيضاً بطريقة المجموعات المتضادة، حيث قورن الأداء الاختبارى لعينة من المتخلفين عقليا والعاديين المتساوين لهم فى العمر الزمنى، ووجدت فروق دالة بينهما لصالح العاديين، وحسب الباحث للصورة المقترحة معاييرها فى صورة نسب ذكاء انحرافية.

أما طبعة ١٩٨٥ من المقياس فلا تكاد توجد دراسة عربية واحدة أجريت عليها.

وقد قام أمين صبرى نور الدين بدراسة استهدفت الكشف عن الأبعاد السيكومترية للتعديل الرابع لمقياس استنفرد – بينيه للذكاء والذى قام بتقنينه كل من ثورنديك وهاجن وستلر سنة ١٩٨٦ (١١٢) (Thorndike, Hagen, & Sattler) حيث قام

الباحث بتعريبه وإعداده للتطبيق لاستخدامه فى البيئة المصرية، وذلك على عينة من أطفال ما قبل المدرسة تألفت من مائتى طفل. وقد أسفرت نتائج الدراسة عن قيم لمعاملات ثبات للدرجة الإجمالية تراوحت من (٠,٨٨) إلى (٠,٩٤). كما حسبت أيضا معاملات الثبات لكل اختبار فرعى من الاختبارات الثمانية لكل فئة عمرية بالإضافة إلى معاملات ثبات المجالات الأربعة للمقياس فقد تراوحت بين (٠,٥٨) إلى (٠,٦٧) مشيرة إلى أن المقياس يقيس الذكاء العام بدرجة مقبولة. كذلك كشفت التحليلات العاملية التوكيدية التى أجراها الباحث من خلال بيانات عينة الدراسة إلى رفض النموذج الهرمى ذى المجالات الأربع كما افترضه مؤلفو المقياس. ومن ثم قام الباحث بمطابقة تلك البيانات مع عدة نماذج أخرى مختلفة، انتهت إلى أن أفضل نموذج يطابق البيانات فيما يتعلق بالأطفال الصغار هو النموذج الثنائى، والذى يفترض أن العامل العام ينبثق منه عاملين هما: الفهم اللفظى والاستدلال غير اللفظى/ البصرى.

أظهرت نتائج الدراسة أيضا أن الدرجات الخام تزيد بتزايد العمر بصورة مطردة وبفروق دالة للأعمار من سن سنتين حتى خمس سنوات فى كل اختبار فرعى مشيرة إلى صدق الاختبار. كما قدم الباحث معايير المقياس والتى تمثلت فى تقديم: جداول تحويل الدرجات الخام لكل اختبار فرعى إلى معايير بمتوسط (٥٠) وانحراف معيارى مقداره (٨) لكل فئة عمرية من الفئات الأربع، وجداول معايير لكل مجال من المجالات الأربعة وللدرجة الإجمالية بمتوسط (١٠٠) وانحراف معيارى (١٦) . علاوة على معايير لاستخراج الدرجة الإجمالية من أى عدد من المجالات الأربعة. وقد كانت جميع خطوات حساب المعايير موافقة لنفس الإجراءات التى اتبعها مؤلفوه، وطبقا لما هى مقررة فى الدليل الفنى للمقياس.

كما قام بتقديم الباحث عدة جداول تفسيرية لتحليل أداءات المفحوص على الاختبارات المختلفة منها: حدود الثقة لدرجات المجالات والدرجة الإجمالية بمستويات دلالة مختلفة لكل فئة عمرية. قدم الباحث كذلك جداول للكشف عن الفروق الدالة بين درجات الاختبارات الفرعية وبين المجالات الأربعة والدرجات الإجمالية عند مستوى ثقة (٠,٠٥) ، (٠,٠١) لكل فئة عمرية، ثم متوسط تلك الفروق. ونظراً لما كشفت عنه نتائج التحليل العاملى من وجود عاملين فقط فقد قدم

الباحث أيضا جدولا للفروق الدالة بين العاملين لكل فئة عمرية. ثم عرض الباحث بعد ذلك إمكانية مقياس استنفرد – بينيه: الرابع فى تحديد المتفوقين والمتخلفين عقليا من واقع الدرجات المعيارية لعينة الدراسة وبعض الاعتبارات الكلينيكية المتعلقة بتطبيق المقياس.

الطبعة الخامسة لمقياس ستانفورد – بينيه*

وظهر مؤخرا الإصدار الخامس من مقياس استنفرد – بينيه للذكاء سنة ٢٠٠٣ Stanford- Binet Intelligence Scales, Fifth Edition (SB5) Galr H. Roid من جامعة واشنطن (Roid, 2003 a, b) (٩٥-٩٦).

ويقدم مقياس ستانفورد – بينيه للذكاء: الإصدار الخامس درجات عشر اختبارات بمتوسط ١٠ وانحراف معيارى ٣ ومدى للدرجات من ١-١٩ والذى يشكل معا بأساليب مختلفة قياس خمسة عوامل، ومن ثم الدرجة الإجمالية للذكاء العام. والعوامل التى يقيسها الإصدار الخامس هى نفس العوامل التى استند عليها ستانفورد – بينيه للذكاء : الإصدار الرابع Thorndike, Hagen, & Sattler, 1986 وهى الإستدلال السائل والمعرفة والإستدلال الكمى والتجهيز البصرى المكانى، وهذه العوامل مشتقة فى الأصل من دراسات كل من Cattel (1971)، وHorn (1994) وCarroll (1993) فى الذكاء غير انه تمت إضافة بعد خامس وهو الذاكرة العاملة Working Memory حيث يشكل كل اختبارين فرعيين من الاختبارات العشرة أحد العوامل الخمسة. وأحد هذان الاختباران يمثل اختبارا لفظيا والثانى غير لفظى. ولذلك فهناك درجة إجمالية تمثل العامل العام وخمس درجات تمثل خمسة عوامل وهناك كذلك درجتان لنسبة الذكاء غير اللفظية Nonverbal IQ ونسبة الذكاء اللفظية Verbal IQ ويستغرق التطبيق من ٤٥ إلى ٧٥ دقيقة Becker,2003 Mleko, & Burns, 2005, Ruf, 2003 .

ويتميز مقياس ستانفورد – بينيه للذكاء : الخامس بما يلى:

– أكثر جاذبية من الإصدارات السابقة من حيث المواد الملونة واللعب والأدوات.

– تم تقنينه وفقا للتعداد السكانى لسنة ٢٠٠٠ بالولايات المتحدة الأمريكية.

* يشكر مؤلفو الكتاب د. أمين صبرى نور الدين على امدادهم بمعلومات عن الطبعة الخامسة لمقياس ستانفورد – بينية.

– تضمن اختبارين لتحديد المسار routing test أحدهما لفظى والآخر غير لفظى.

– يقدم كذلك درجات للعوامل المقاسة علاوة على الدرجة الإجمالية كما هو الحال فى الإصدار الرابع ١٩٨٦ .

– يحتوى على مفردات مشتركة مع الاصدارات السابقة للتواصل مع تلك الإصدارات.

– يغطى مدى عمرى أوسع من أى إصدار سابق (من سن سنتين وحتى فوق ٨٥ سنة).

– يقدم درجة نسبة ذكاء غير لفظية متلافيا بذلك أوجه النقد التى صوبت إلى الإصدارات السابقة.

– الدرجة الإجمالية متوسطها ١٠٠ وانحرافها المعيارى ١٥ بدلا من ١٦ كما كان فى الإصدارات السابقة لكى يسهل مقارنته مع المقاييس الأخرى. وفيما يتعلق بالاختبارات الفرعية فقد كان متوسطها ١٠ وانحرافها المعيارى ٣.

– يستخدم مقياس ستانفورد – بينيه للذكاء الإصدار الخامس الاختبار التكيفى adaptive testing لتسهيل عملية التطبيق وتقليل جهد المفحوص.

– له إطار نظرى واضح وله سقف مفردات عال للتعرف على المتفوقين عقليا وكذا أرضية مفردات منخفضة للتعرف المبكر على المعاقين واصحاب الصعوبات المعرفية.

– اشتمل على محتوى متوازن ومتساو من الاختبارات اللفظية وغير اللفظية فى جميع العوامل بخلاف الإصدارات السابقة التى كانت متشبعة بالعامل اللفظى. ويبين الجدول رقم (٢٣) نسب محتويات المفردات عبر الإصدارات المختلفة.

جدول رقم (٢٣)

يبين نسب محتويات المفردات فى الإصدارات المختلفة

العامل / الإصدار	س ب ٥	س ب ٤	الصورة ل – م	الصورة ل	الصورة م	١٩١٦
المعرفة	٢٦٪	٢٧٪	٤٠٪	٤١٪	٣٨٪	٣٨٪
الاستدلال السائل	١٧٪	١٤٪	٢٠٪	١٧٪	٢١٪	١٧٪
التجهيز البصرى – المكانى	١٨٪	١٧٪	١٨٪	١٦٪	١٥٪	١١٪
الاستدلال الكمى	٢١٪	١٧٪	٦٪	٦٪	٦٪	٨٪
الذاكرة العاملة	١٢٪	٢٪	٤٪	٣٪	٤٪	٩٪
الذاكرة قصيرة المدى	٦٪	٢٢٪	٩٪	١٣٪	١١٪	١١٪
أخرى **	٠٪	٠٪	٤٪	٣٪	٥٪	٦٪

س ب = ستانفورد – بينيه للذكاء

* العلاقات اللفظية حسبت داخل الاستدلال السائل بالإضافة إلى المعرفة. سلاسل الأرقام حسبت فى الاستدلال السائل بالاضافة إلى الاستدلال الكمى.

** هذه المفردات تغطى أى قدرات تمثل التجهيز السمعى والاسترجاع طويل المدى وتقديرات الأوزان.

ويؤكد Becker, 2003; Roid, & Carson 2003 قدرة ستانفورد – بينيه للذكاء: الإصدار الخامس على تشخيص والتنبؤ بالأطفال ذوى صعوبات التعلم واضطرابات الانتباه والمتخلفين عقليا وبطيئى التعلم. غير أن (Minton, & Pratt, 2006) يحذران من أن ستانفورد – بينيه للذكاء: الإصدار الخامس يعطى درجات أقل بدلالة من مقياس وكسلر للذكاء: الثالث وذلك للأطفال المتفوقين عقليا.

مقاييس وكسلر للذكاء

ظهرت الطبعة الأولى من مقاييس وكسلر عام ١٩٣٩ وعرفت حينئذ باسم مقياس وكسلر- بليفو للذكاء Wechaler - Bellevue، وكان أحد الأغراض المهمة من إعداد هذه المقاييس توفير اختبارللذكاء يصلح للراشدين وقد أشار وكسلر فى ذلك الحين إلى أن الاختبارات التى كانت متاحة حتى ذلك الوقت (ومنها مقياس ستانفورد- بينيه) معدة فى جوهرها للأطفال وتلاميذ المدارس، وكانت تستخدم مع المراهقين والراشدين عن طريق إضافة بعض الأسئلة الأكثر صعوبة من نفس النوع، ولذلك فإن من المشكلات المهمة فى هذه الاختبارات أن محتواها لا يثير اهتمام الراشدين (أى يعوزها الصدق الظاهرى) . وإذا لم يتوافر فى المقياس حدا أدنى من الصدق الظاهرى يصعب تكوين علاقة طيبة بين الفاحص والمفحوص الراشد.

وبالإضافة إلى ذلك فإن المبالغة فى تأكيد عامل السرعة فى معظم الاختبارات الفردية يؤدى إلى تعطيل أداء الراشدين، وكذلك فإن الاهتمام بالتناول الروتينى للنواحى اللفظية تحظى بالاهتمام الأكبر فى هذه المقاييس، ثم إنه يستحيل استخدام معيار العمر العقلى مع الراشدين كما أشرنا، وأخيراً فإن عينات تقنين الاختبارات الفردية (مثل مقياس ستانفورد- بينيه) لم تتضمن إلا قليلا من الراشدين.

ولمواجهة هذه المشكلات أعد وكسلر مقاييسه للذكاء، ومن أشهرها مقياسه لذكاء الراشدين، والذى يسمي فى الوقت الحاضر Wechsler Adult Intelligence Scale WAIS ، والذى ظهر فى عام ١٩٥٥ ليحل محل مقياس وكسلر- بلفيو القديم، وقد ظهرت من المقياس طبعة جديدة معدلة فى أواخر السبعينات باسم WAIS-R لقياس ذكاء الراشدين من سن ١٦ إلى سن ٧٤ سنة، وبالطبع يتشابه المقياس الجديد فى الشكل والمحتوى مع المقياس السابق، إلا أنه تغلب على بعض الصعوبات الفنية الخاصة بمدى تمثيل عينة التقنين، وثبات الاختبارات الفرعية التى يتكون منها المقياس، ثم ظهرت منه عدة طبعات آخرها ما صدر عام ١٩٨١ باستخدام مقياس وكسلر لذكاء الراشدين المعدل WAIS-R.

وقد ظهر مقياس وكسلر- بلفيو الأصلى باللغة العربية فى عام ١٩٥٦ وأعده لويس كامل مليكة ومحمد عماد الدين إسماعيل (٤٨) ومنذ ذلك الحين لم تطرأ عليه تعديلات تذكر.

وفى عام ١٩٤٦ نشر وكسلر صورة ثانية لمقياس بلفيو الأصلى قننت على ألف راشد من الذكور أعمارهم تمتد من ١٨ – ٤٠ سنة، إلا أن هذه الصورة سرعان ما فقدت هويتها لأنها صارت أساسا لبناء مقياس آخر لذكاء الأطفال الذى شاع باسم مقياس وكسلر لذكاء الأطفال (WISC) والذى ظهرت منه طبعة معدلة عام ١٩٧٤ باسم (WISC-R) لقياس ذكاء الأطفال من سن ٦ إلى سن ١٦ سنة.

وقد نقل المقياس الأصلى إلى اللغة العربية محمد عماد الدين إسماعيل ولويس كامل مليكه عام ١٩٦١ ولا تتوافر عنه بيانات سيكومترية فى اللغة العربية حتى الآن.

وسرعان ما ظهرت الحاجة إلى قياس ذكاء الأطفال فى سن ما قبل السادسة فأصدر وكسلر فى عام ١٩٦٧ اختباره الثالث والمهم والذى عرف باسم مقياس وكسلر لذكاء أطفال ما قبل المدرسة والذى عرف باسم (WPPSI) ويقيس ذكاء الأطفال بين ٤، ٦,٥ سنة.

وتتشابه المقاييس الثلاثة فى البنية العامة حيث يتضمن المقياس مجموعة من الاختبارات الفرعية تشتق منها درجة لفظية Verbal ومجموعة أخرى تشتق منها درجة أدائية performance ، وترتبط الدرجتان معا فى درجة كلية Full Scale. وتتشابه الاختبارات الفرعية فى المستويات العمرية المختلفة إلا أنها لا تتطابق فهى تختلف فى المحتوى والطول ومستويات الصعوبة والوزن النسبى للسؤال، ووجه التطابق الوحيد هو المعيارالمستخدم، ففى الاختبارات الفرعية تحول الدرجات الخام إلى درجات معيارية معدلة متوسطها ١٠ وانحرافها المعيارى ٣، أما فى الاختبار الكلى فإن الدرجة المعيارية المعدلة متوسطها ١٠٠ وانحرافها المعيارى ١٥.

ويوضح الجدول رقم (٢٤) الاختبارات الفرعية التى تتألف منها المقاييس الثلاثة ويمكن الاستفادة به فى المقارنة.

جدول رقم (٢٤) الاختبارات الفرعية فى مقاييس وكسلر الثلاثة

WAIS-R مقياس الراشدين	WISC-R مقياس الأطفال	WPPSI-R مقياس ما قبل المدرسة
المقياس اللفظى		
المعلومات العامة	المعلومات العامة	المعلومات العامة
الفهم العام	الفهم العام	الفهم العام
الاستدلال الحسابى	الاستدلال الحسابى	الاستدلال الحسابى
المتشابهات	المتشابهات	المتشابهات
المفردات	المفردات	المفردات
مدى الأرقام	(مدى الأرقام)	(الجمل)
المقياس الأدائى		
رسوم المكعبات	رسوم المكعبات	رسوم المكعبات
تكميل الصور	تكميل الصور	تكميل الصور
ترتيب الصور	ترتيب الصور	--
تجميع الأشياء	تجميع الأشياء	--
رموز الأرقام	الترميز	بيت الحيوان
--	(المتاهات)	المتاهات
--	--	الرسوم الهندسية

(١) الاختبارات بين القوسين هى اختبارات احتياطية.

ويلاحظ أن عدد الاختبارات فى كل فئة فرعية لا يتجاوز سبع اختبارات ولا تقل عن خمسة اختبارات. ويمكن أن يستخرج من كل فئة (لفظية أو أدائية) نسبة ذكاء مستقلة غير تلك المستمدة من المقياس ككل.

وعند البدء عادة فى تطبيق المقياس، لأى فئة عمرية، يبدأ الفاحص باختبار المعلومات، ويلاحظ أنه لا يتضمن أى معلومات تخصصية، وأن الأسئلة المتضمنة فى

بداية الاختبار سهلة ويمكن اجتيازها بسهولة من معظم المفحوصين فيما عدا أولئك الذى يعانون من بعض المشكلات (التخلف العقلى مثلا) وفى ضوء أدائهم يقرر الفاحص ما إذا كان سيستمر فى تطبيق باقى الإختبارات أو التوقف.

ونعرض فيما يلى وصفا مفصلا لكل مقياس من المقاييس الثلاثة، ونبدأ بمقياس الراشدين ثم مقياس الأطفال، ثم مقياس أطفال مرحلة ما قبل المدرسة.

وصف مقياس وكسلر لذكاء الراشدين: (WAIS-R)*

يتكون مقياس WAIS-R* من ١١ اختبارا فرعيا، ٦ منها تصنف فى مقياس لفظى verbal scale ، ٥ منها تصنف فى مقياس أدائى أو عملى preformance، وفيما يلى وصف موجز لهذه الاختبارات الفرعية مع بيان عدد الأسئلة التى يتكون منها كل اختبار فى الطبعتين العربية والأجنبية.

المقياس اللفظى: يتكون المقياس اللفظى من ٦ اختبارات هى:

١ – المعلومات العامة: information، ويتكون من ٢٩ سؤالا فى الطبعة الأجنبية (٢٦ سؤالا فى الطبعة العربية منها سؤال تمهيدى) تشمل مدى واسعا من المعلومات العامة التى يفترض شيوعها فى ثقافة الراشدين، وقد روعى تحرر الاختبار من المعارف التخصصية والأكاديمية.

٢ – الفهم العام: comprehension، ويتكون من ١٤ سؤالا فى الطبعة الأجنبية (١٠ أسئلة فى الطبعة العربية)، وفى كل سؤال يشرح للمفحوص ما يجب عمله فى ظروف معينة، ويتطلب هذا الاختبار قدرة على الحكم العملى والفهم العام Common Sense، وهو يشبه فى هذا أسئلة الفهم فى مقياس ستانفورد ـ بينيه ، إلا أن محتواه أقرب إلى اهتمامات الراشدين ونشاطهم.

٣ – الاستدلال الحسابى arithmetic ويتكون من ١٤ مسألة حسابية فى الطبعة الأجنبية (١٠ مسائل فى الطبعة العربية) تتفق مع مستوى مادة الحساب فى المرحلة الابتدائية، وتعرض كل مسألة شفويا، ويحلها المفحوص شفويا كذلك دون استخدام الورقة والقلم، ولكل مسألة زمن محدد.

---

* WAIS - R = Wechsler Adult Intelligence Scale - Revised.

٤- المتشابهات: Similarities، ويتكون من ١٣ سؤالا فى الطبعة الأجنبية (١٣ سؤالا فى الطبعة العربية) تطلب من المفحوص أن يحدد كيف يتشابه شيئان (البرتقالة والموزة مثلا).

٥ - مدى الأرقام (أو إعادة الأرقام): digit span، وفيه تعرض على المفحوص قوائم تتكون من ٣ أرقام إلى ٩ أرقام، ويكون عليه إعادتها شفويا، وفى القسم الثانى من الاختبار يكون على المفحوص إعادة الأرقام بالعكس.

٦ - المفردات : vocabulary ، وفيه تعرض على المفحوص شفويا وبصريا فى وقت واحد ٤٢ كلمة متزايدة الصعوبة، ويطلب منه تحديد معنى كل كلمة.

المقياس الأدائى أو العملى: يتكون المقياس العملى من ٥ اختبارات هى:

٧ - رموز الأرقام Digit symbol وهو عبارة عن صورة من اختبار التعويض الشفرى Code substitution المألوف فى مقاييس الذكاء غير اللغوية، ويتكون مفتاح الشفرة من ٩ رموز تتزاوج مع الأرقام التسعة. ويعطى هذا المفتاح للمفحوص ويكون عليه أن يكتب أكبر عدد من الرموز التى تدل على الأعداد الموجودة فى ورقة الإجابة فى زمن لايتجاوز ١,٥ دقيقة.

٨ - تكميل الصور: Picture completion ويتكون من ٢١ بطاقة فى الطبعة الأجنبية (١٥ فى الطبعة العربية) كل منها يحتوى على صورة ناقصة، ويكون على المفحوص أن يذكر الجزء الناقص من كل صورة.

٩ - رسوم المكعبات: block design وتشبه مواد هذا الاختبار ما يوجد فى اختبار Kohs Bloch Design Test ويتكون من صندوق به ١٦ مكعبا، و٩ بطاقات على كل منها رسم مختلف (اثنتان من هذه البطاقات للتدريب، وقد لون أوجه المكعبات بالألوان الآتية فى الأصل الأجنبى: الأحمر فقط والأبيض والأحمر معا، أما فى الطبعة العربية فهى ملونة بالأزرق فقط والأبيض فقط والأصفر فقط والأبيض والأحمر معا والأزرق والأصفر معا، ويطلب من المفحوص إنتاج رسوم على النحو الذى تحدده البطاقات ذات مستويات متزايدة من التعقد ما بين ٤ مكعبات، ٩ مكعبات، ولكل بطاقة زمن محدد.

١٠ – ترتيب الصور: picture arrangement، ويتكون كل سؤال فيه من مجموعة من البطاقات، عليها صور، ويطلب من المفحوص ترتيبها بالتتابع الصحيح بحيث تعطى قصة مفهومة، ولكل سؤال وقت محدد (عدد الأسئلة فى الطبعة الأجنبية ٨ وفى الطبعة العربية ٦).

١١ – تجميع الأشياء Object assembly ، ويتبع هذا الاختبار نفس النمط المستخدم فى اختبار Pinter-Paterson الأدائى، ويتكون فى الطبعة العربية من نماذج من الخشب لثلاثة أشياء هى الصبى أو المانيكان، والوجه أوالبروفيل، واليد وقد قطع كل منها إلى قطع مختلفة، ويطلب من المفحوص فى كل منها جمع القطع بحيث تكون الشكل الكامل.

تصحيح المقياس:

تصحح اختبارات الاستدلال الحسابى ورسوم المكعبات وترتيب الصور وتجميع الأشياء لسرعة الأداء ودقته على النحو التالى:

١ – الاستدلال الحسابى: تعطى درجة واحدة لكل مسألة تحل حلا صحيحا فى الوقت المحدد لها ، وتعطى درجة واحدة إضَافية للزمن فى المسألتين ٩، ١٠ إذا حلت كل منهما فى حدود ٤٠ ثانية (الزمن العادى لكل منهما ١٢٠ ثانية)، وتعطى درجتان إضافيتان إذا حلت المسألة فى حدود ١٥ ثانية، يحسب الزمن بعد انتهاء المفحوص من قراءة المسألة (النهاية العظمى ١٤).

٢ – رسوم المكعبات: تقدر الرسوم حسب الدقة والزمن الأقصى، وتعطى درجات إضافية فى حالة إتمام الرسم فى زمن أقل، وقد أعد جدول لإعطاء تقديرات حسب زمن الإنجاز (النهاية العظمى ٤٢).

٣ – ترتيب الصور: تقدر الدرجة فى هذا الاختبار على أساس دقة الترتيب والزمن المستغرق فيه، وأعدت جداول لهذا الغرض بحيث تصحح كل من المجموعات الثلاثة الأولى على أساس الكل أو لا شىء (صواب أو خطأ) أما المجموعات الثلاث الأخيرة فتعطى تقديرات جزئية على كل ترتيب لايطابق الترتيب الصحيح، ولكن يمكن قبوله، وبالنسبة للمجموعتين الأخيرتين تعطى تقديرات إضافية على الترتيب الصحيح إذا تم فى زمن معين، وتحسب الدرجة النهائية فى هذا الاختبار بجمع التقديرات الجزئية فى المجموعات المختلفة (والنهاية العظمى ٢١).

٤ – تجميع الأشياء: يصحح الشكل الأول (الصبى) حسب الدقة فقط، أما الشكلان الآخران (الوجه واليد) فيصحح كل منهما حسب الدقة والزمن معا. وفى الحالة الأخيرة تكون الدرجة فى مواد الاختبار هى المجموع الكلى لتقديرات الدقة والزمن جميعا، وتوجد بكراسة التعليمات بروتوكولات خاصة لتصحيح كل شكل حسب الدقة، وجدول للتصحيح على أساس الزمن (النهاية العظمى ٢٦).

أما اختبارات المعلومات العامة، ومدى الأرقام (أو إعادة الأرقام)، وتكميل الصور، ورموز الأرقام فإن الدرجة الكلية فيها هى مجموع الإجابات الصحيحة حسب مفتاح التصحيح المبين فى كراسة التعليمات، والنهايات العظمى لهذه الاختبارات الأربع على التوالى هى: ٢٥، ١٧، ١٥، ٦٧.

وقد قام لويس كامل مليكه (٤٢) بإعداد نماذج تصحيح خاصة باختبارات الفهم العام، والمتشابهات، والمفردات، نشرها فى كراسة منفصلة عام ١٩٦٠، وفى هذه النماذج تعطى لأسئلة اختبارى الفهم العام والمتشابهات درجتان أودرجة واحدة أو صفر حسب درجة التعميم فى الإجابة ونوعه وبهذا تكون النهاية العظمى لكل من هذين الاختبارين هى: ٢٠، ٢٤ على التوالى، أما اختبار المفردات فإن أسئلته تعطى فى نماذج التصحيح درجة واحدة أونصف درجة أو صفر (النهاية العظمى ٢٤).

المعايير:

لقد كانت عينة التقنين المستخدمة فى مقياس وكسلر للراشدين فى طبعاته المختلفة من العينات القلائل التى يتوفر فيها شرط التمثيل، فقد كانت تتألف من ١٧٠٠ شخصاً فى المقياس الأصلى، أما المقياس المعدل الذى صدر عام ١٩٨١ فقد اعتمد على عينة تقنين قوامها ١٨٨٠ تتساوى فيها أعداد الرجال والنساء موزعين فى مستويات عمرية سبع بين ١٦ سنة و٧٤ سنة اختيرا بحيث تتوافر بهم خصائص التعداد العام للسكان فى الولايات المتحدة الذى أجرى عام ١٩٧٠ من حيث المنطقة الجغرافية والبيئة الريفية والحضرية والأصل العنصرى (وخاصة البيض والزنوج)، والمستوى المهنى، والمستوى التعليمى، واقتصرت العينة على «الراشدين الأسوياء»، واستبعدت حالات التخلف العقلى، والمصابون بإصابات المخ، والذين يعانون من اضرابات سلوكية وانفعالية أونقائص جسمية، كما استبعدت الحالات التى لا تستطيع استخدام اللغة الإنجليزية.

وقد حولت الدرجات الخام فى الاختبارات الفرعية إلى درجات معيارية معدلة متوسطها ١٠ وانحرافها المعيارى ٣، وكانت هذه الدرجات الموزونة Scaled Score مشتقة من جماعة مرجعية عددها ٥٠٠ حالة شملت جميع المفحوصين من سن ٢٠ سنة ، ٣٤ سنة فى عينة التقنين، وهكذا أصبحت الدرجات الفرعية فى صورة وحدات قابلة للمقارنة، ثم حسبت درجات المقياس اللفظى،والمقياس الأدائى، والمقياس الكلى بجمع الدرجات الموزونة فى الاختباراتِ اللفظية الست،والاختبارات الأدائية الخمس، والاختبارات الأحد عشر جميعا على التوالى، ثم حولت هذه الدرجات الثلاثة إلى نسبة ذكاء انحرافية متوسطها ١٠٠ وانحرافها المعيارى ١٥ فى كل مجموعة عمرية، وهكذا يتحدد الوضع النسبى للمفحوص بمقارنته بالأشخاص من نفس مستواه العمرى.

وقد استخدمت نفس الإجراءات فى تقنين الطبعة العربية حيث استخدمت فى إعداد الدرجات الموزونة ثلاث فئات عمرية هى:

١ – من ٢٠ إلى أقل من ٢٥ سنة وعدد أفرادها ١١٤ .
٢ – من ٢٥ إلى أقل من ٣٠ سنة وعدد أفرادها ١١٢ .
٣ – من ٣٠ إلى أقل من ٣٥ سنة وعدد أفرادها ١٠٢ .

وقد استخدمت هذه العينة البالغ عددها ٣٢٨ مفحوصاً فى إعداد نماذج ومعايير التصحيح لاختبارات الفهم والمتشابهات والمفردات بالإضافة إلى إعداد جدول الدرجات الموزونة وجداول لنسب الذكاء الانحرافية لكل فئة من فئات السن (٤٢) .

وفى تصحيح هذا المقياس تعطى للمفحوص درجات ثلاث: نسبة الذكاء الكلية ونسبة الذكاء اللفظية ونسبة الذكاء العملية، ويعطى فى الطبعة الأجنبية درجة رابعة تدل على ما يسميه وكسلر «معامل الكفاءة» efficiency quotient أو (E Q)، ويتحدد بتقدير نسبة ذكاء الفرد فى الفئة من ٢٠ إلى ٢٣ (فى اختبار وكسلر بلفيو طبعة ١٩٣٩) وفى الفئة ٢٥ – ٢٩ (فى اختبار وكسلر لذكاء الراشدين طبعة ١٩٥٥)، بصرف النظر عن العمر الحقيقى للمفحوص، ويعتمد هذا المعامل على أساس أن الفئة من ٢٠ – ٢٤ (فى تقنين (١٩٣٩) والفئة من ٢٥ – ٢٩ (فى تقنين ١٩٥٥) هما اللتان وصلتا إلى أعلى مستويات الأداء فى المقياس، ومعنى ذلك أن معامل الكفاءة طريقة للحكم على أداء الفرد فى ضوء متوسط أداء الأفراد من المستوى العمرى الذى يصل فيه الأداء الاختبارى أقصاه .

ويذكر لويس كامل مليكه (٤٨ : ٦١) أنه يمكن الحصول على معامل الكفاءة بصورة مؤقتة عن طريق استخدام جداول نسب ذكاء الفئة من ٢٠ – ٢٤ من العينة المصرية آنفة الذكر، لأن هذه الفئة حققت بالفعل أعلى مستوى من الأداء فى المقياس.

تفسير المعايير:

أعد وكسلر عام ١٩٤٤ فئات الذكاء التى يمكن أن تستخدم فى أغراض التصنيف الكلينيكى. وكان أساس التصنيف هو حدود الخطأ المعيارى للمتوسط، فالشخص الذى تكون درجته أقل من المتوسط بمقدار يبلغ ثلاثة أمثال الخطأ أو أكثر يعد من فئة ضعاف العقل، أما الذى يحصل على درجة تقع ما بين خطأين معياريين وثلاثة أخطاء معيارية فيعد من قبيل الحالات الهامشية، وهكذا، ويبين الجدول رقم (٢٥) أسس تصنيف مستويات الذكاء فى مقياس وكسلر.

جدول رقم (٢٥) تصنيف مستويات الذكاء فى مقياس وكسلر

التصنيف	حدود الخطأ المعيارى	حدود نسبة الذكاء الانحرافية	النسبة المئوية فى الأصل الإحصائى
ضعاف العقول	من ٣ فأقل	٦٥ فأقل	٢,٢
الحالات الهامشية	من – ٢ حتى – ٣	من ٦٦ حتى ٧٩	٦,٧
الأغبياء	من – ١ حتى – ٢	من ٨٠ حتى ٩٠	١٦,١
المتوسطون	من – ١ حتى + ١	من ٩١ حتى ١١٠	٥٠,٠
أعلى من المتوسط	من + ١ حتى + ٢	من ١١١ حتى ١١٩	١٦,١
الممتازون	من + ٢ حتى + ٣	من ١٢٠ حتى ١٢٧	٦,٧
الممتازون جداً	من + ٣ فأكثر	من ١٢٨ فأكثر	٢,٢

ثبات المقياس:

حسبت معاملات الثبات للطبعة الأجنبية من مقياس وكسلر لذكاء الراشدين لمجموعات من فئات الأعمار ١٨ – ١٩، ٢٤ – ٣٤، ٤٥ – ٥٤ اختيرت من عينة التقنين على أساس تمثيلها لمدى العمر الذى تشمله هذه العينة، واستخدمت طريقة التجزئة النصفية (بعد تصحيح سبيرمان - براون) فى حساب ثبات كل اختبار فرعى

فيما عدا اختبارى رموز الأرقام ومدى (إعادة) الأرقام حيث استخدمت معهما طريقة الصور المتكافئة لعدم صلاحية طريقة التجزئة النصفية لهما بسبب غلبة عامل السرعة الشديدة عليهما.

ويلاحظ أن معاملات الثبات للمقياس الكلى قد بلغت فى المجموعات الثلاثة ٠,٩٧، وبلغت فى المقياس اللفظى ٠,٩٦، وفى المقياس الأدائى ٠,٩٤ ومعنى هذا أن المقاييس الثلاثة على درجة عالية من الثبات فى ضوء الاتساق الداخلى.

أما فى الطبعة العربية فقد حسب فرج عبدالقادر طه (٣٧ : ٧٨) ثبات المقياس بطريقة إعادة الاختبار وطريقة التجزئة النصفية على عينة من عمال الصناعة، وفى الطريقة الأولى أعيد تطبيق المقياس على ٤٠ حالة من العمال فتراوحت معاملات الثبات بين ٠,٦، ٠,٩ بالنسبة لكل من نسب الذكاء والاختبارات الفرعية، أما طريقة التجزئة النصفية فقد حسبت لدرجات ٧٠ عاملا وتراوحت معاملات الثبات الناتجة عنها بين ٠,٥، ٠,٩ لنسبة الذكاء والاختبارات الفرعية، ولم يوضح الباحث كيف استخدم طريقة التجزئة النصفية لاختبارى رموز الأرقام وإعادة الأرقام بالرغم من عدم صلاحية هذه الطريقة لهما كما أشارت نتائج بحوث التقنين التى أجريت على المقياس (٦١ : ٢١٩).

صدق المقياس:

يفضل وكسلر (١١٥) فى تناوله لصدق مقياسه أن يؤكد «صدق المحتوى». ومن رأيه أن الاختبارات الأحد عشر المتضمنة فيه تتفق مع تعريف الذكاء، وأن اختبارات مماثلة استخدمت فى مقاييس الذكاء السابقة، وثبت جدواها فى الأغراض الكلينيكية وهذا القول أقرب إلى صدق التكوين الفرضى ولكن بدون شواهد امبريقية، وبالإضافة إلى ذلك يذكر وكسلر (١١٥) وماتارازو (٨٧) بعض البيانات عن الصدق التلازمى للمقياس، ومن ذلك المقارنة بين المجموعات التعليمية والمهنية المختلفة، بالإضافة إلى حساب معاملات الارتباط مع تقديرات الأداء على العمل والتقديرات المدرسية، وكانت أغلب الفروق بين المجموعات فى الاتجاه المتوقع، ومن ذلك أن درجات العمال أعلى فى الاختبارات العملية، وأن درجات الأفراد فى المهن غير اليدوية أعلى فى الاختبارات اللفظية، أما عن معاملات الارتباط بالمحكات المهنية والمدرسية فقد أعطى المقياس اللفظى قيما أعلى من المقياس الكلى أما المقياس الأدائى فقد أعطى أدنى القيم، وفى كل الأحوال كانت معاملات ارتباط المقياس اللفظى أدنى من معاملات مقياس ستانفورد-

بينيه، أما عن الصدق التنبؤى للمقياس فقد ثبت أن المقياس يتنبأ بدرجة مقبولة بالضعف العقلى كما تؤكد بحوث جيرتن وزملائه (٧٩، ٨٠).

وقد حسبت معاملات الارتباط بين مقياس وكسلر وغيره من الاختبارات والمقاييس المعدة لقياس الذكاء، وكانت النتائج كما يلى:

١ – معامل الارتباط بين المقياس ومقياس ستانفورد ـ بينيه ٠,٨٠ وهو أعلى فى المقياس اللفظى منه فى المقياس الكلى والمقياس الأدائى.

٢ – معاملات الارتباط بين المقياس واختبارات الذكاء الجماعية تتراوح بين ٠,٤٠، ٠,٨٠.

٣ – معامل الارتباط بين المقياس الأدائى واختبارمنيسوتا للوحدات والأشكال ٠,٧٢.

٤ – معامل الارتباط بين المقياس الأدائى واختبار رافن للمصفوفات المتتابعة ٠,٧٠.

٥ – معامل الارتباط بين المقياس الأدائى واختبار بينت للفهم الميكانيكى ٠,٧٠.

صدق التكوين الفرضى:

أجريت عدة تحليلات عاملية للاختبارات الفرعية التى يتألف منها مقياس وكسلر، ومن أشهر هذه الدراسات بحوث كوهن (٧٠، ٧١) على مصفوفات ارتباطية لمجموعات عمرية أربعة من عينة التقنين هى: (من ١٨ – ١٩، ومن ٢٥ – ٣٤، ومن ٤٥ – ٥٤، ومن ٦٠ – ٧٥ +)، وقد وجد الباحث أن الاختبارات الفرعية جميعا تشترك فى عامل واحد يسهم بحوالى ٥٠٪ من التباين الكلى للبطارية، وبالإضافة إلى ذلك استخراج ثلاثة عوامل طائفية هى:

١ – عامل الفهم اللفظى وتشبع باختبارات المفردات والمعلومات العامة والفهم العام والمتشابهات.

٢ – عامل التنظيم الإدراكى وتشبع باختبارات رسوم المكعبات وتجميع الأشياء، وهو عبارة عن مركب من عامل السرعة الإدراكية والتصور البصرى المكانى اللذين أكدتهما البحوث العاملية الأساسية وبخاصة بحوث ثرستون.

٣ – عامل الذاكرة وتشبع باختبارات الاستدلال الحسابى ومدى (إعادة) الأرقام، ويتضمن كلا من الذاكرة الصماء المباشرة للمواد الجديدة واستدعاء المواد التى سبق تعلمها.

ومن الطريف أن نذكر أن نتائج هذا البحث لم يدعم الإجراء الذى يستخدم فى تصحيح هذا المقياس أى تقسيمه إلى مقياس لفظى ومقياس عملى أدائى لكل منهما نسبة ذكاء منفصلة،ورغم أن استخدام نسب الذكاء الكلية يدعمه وجود عامل عام مشترك فى جميع الاختبارات، فإن عامل الفهم اللفظى لم يظهر إلا فى ٤ اختبارات لفظية فقط من ٦ اختبارات، أما عامل الذاكرة فقد وجد فى الاختبارين اللفظيين الباقيين بالإضافة إلى اختبارات فرعية أخرى من المقياسين عند الراشدين الأكبر سنا. أما عامل التنظيم الإدراكى فكانت تشبعاته الدالة فى اختبارين فقط من الاختبارات الأدائية الخمس، وأما الاختبارات الأدائية الثلاث الباقية فهى فيما يبدو لها عواملها الخاصة التى لا تشترك مع الاختبارات الأخرى المتضمنة فى بطارية المقياس.

وقد قام محمد عماد الدين إسماعيل (٥٠) بتحليل عاملى لمقياس وكسلر (الطبعة العربية) توصل منه إلى عامل عام واحد مشترك فى جميع اختبارات المقياس لفئة السن من ١٦ – ١٩، وكانت أكثر الاختبارات تشبعا به اختبار المعلومات العامة ثم اختبار الفهم العام، وتراوحت تشبعات الاختبارات الفرعية به بين ٠,٩٥، ٠,٦٤ ومن الطريف أن الباحث يطابق بين العامل العام والذى وجده فى بطارية وكسلر والعامل العام عند سبيرمان، أما بالنسبة لفئة السن من ٢٠ – ٢٤ فقد وجد عاملين: عامل عام مشترك فى جميع الاختبارات، وعامل ثنائى القطب تشبع به تشبعات موجبة كاختبارات المعلومات والفهم والمتشابهات والمفردات، وتشبع به تشبعات سالبة كاختبارات ترتيب الصور وتكميل الصور وتجميع الأشياء ويمكن اعتباره عاملا يميز بين الاختبارات اللفظية فى مقابل الاختبارات غير اللفظية.

وأخيراً نذكر أن البحوث العاملية العديدة التى قام بها سوندرز (١٠٥، ١٠٦، ١٠٧، ١٠٨)، مستخدما مصفوفات ارتباط الأسئلة الفردية للاختبارات الفرعية للمقياس، توصلت إلى ١٠ عوامل على الأقل، ولم يجد الباحث تطابقا كاملا بينها وبين الاختبارات الفرعية للمقياس، بل تأكد أن بنية بعض الاختبارات الفرعية على درجة كبيرة من

«التعقد العاملى»، وأن بعض العوامل تتعدى أكثر من اختبار فرعى واحد.

ولم يظهر لهذا المقياس صدى يذكر فى البيئات العربية الأخرى بعد ظهور الصورة المصرية للمقياس، باستثناء ما قام به فرج عبدالقادر طه وصلاح أحمد مرحاب (٣٨) بإعدادهما الصورة المغربية للمقياس، وقاما بتقنينها على عينة من ١٥٠ مفحوصًا من الجنسين ومن الأعمار من ٢٠ – ٢٤ سنة، وحسب الثبات عن طريق إعادة الاختبار للاختبارات الأحد عشر وتراوحت معاملات الثبات بين ٠,٧٧ ، ٠,٩٦ وقدر الصدق باستخدام مجموعتين متضادتين من ضعاف العقول والأذكياء وكانت الفروق بينهما دالة فى المقاييس الثلاثة، وأعدت جداول الدرجات الموزونة ونسب الذكاء الثلاثة.

وصف مقياس وكسلر لذكاء الأطفال: (WISC)

يتكون مقياس وكسلر لذكاء الأطفال Wechsler Intelligence Scale for Children أو WISC من الأسئلة السهلة التى كان يتألف منها، مقياس وكسلر ـ بلفيو الأصلى، ويقسم المقياس ـ كما هو الحال فى مقياس الراشدين ـ إلى قسمين : أحدهما لفظى والآخر عملى، والمجموع الكلى للاختبارات الفرعية ١٢ اختبارا منها اختباران احتياطيان أو إضافيان عند الضرورة، على النحو الآتى:

القسم اللفظى	القسم العملى (الأدائى)
(١) المعلومات العامة	(٦) تكملة الصور
(٢) الفهم العام	(٧) ترتيب الصور
(٣) الاستدلال الحسابى	(٨) تصميم (رسم) المكعبات
(٤) المتشابهات	(٩) تجميع الأشياء
(٥) المفردات	(١٠) الترميز
(إعادة الأرقام)	(المتاهات)

وقد اعتبر وكسلر الاختبارات الاحتياطية هى تلك التى ارتبطت أدنى الارتباطات ببقية المقياس، وعلى ذلك فإن اختبار مدى (إعادة) الأرقام فى المقياس اللفظى واختبار المتاهات فى المقياس العملى من هذا النوع. ويتطابق اختبار الترميز مع اختبار رموز الأرقام فى مقياس الراشدين فيما عدا غلبة الأسئلة السهلة عليه. أما

الاختبار الوحيد الذى ليس له نظير فى مقياس الراشدين فهو اختبار المتاهات، ويتكون من ٨ متاهات ورقية متزايدة الصعوبة، وتصحح فى ضوء الزمن وعدد الأخطاء، وفى رأى وكسلر أن اختبارى الترميز والمتاهات يمكن أن يحل أحدهما محل الآخر، ويتطلب اختبار الترميز زمنا أقل من اختبار المتاهات، وقد يفضله بعض الفاحصين لهذا السبب.

تصحيح المقياس وتفسير الدرجات:

لايختلف مقياس الأطفال فى إجراءات تصحيحه من مقياس الراشدين إلا اختلافات طفيفة. وتتلخص هذه الأجراءات فى تحويل الدرجات الخام فى كل اختبار فرعى إلى درجات معيارية معدلة فى إطار المجموعات العمرية التى ينتمى إليها المفحوص. وقد أعد وكسلر جداول لهذه الدرجات الموزونة لفئات عمرية مقدارها أربعة شهور ما بين سن ٥ سنوات حتى سن ١٥ سنة، وتشبه الدرجات الموزونة هنا درجات مقياس الراشدين فى أن متوسطها ١٠ وانحرافها المعيارى ٣، ثم تجمع الدرجات الموزونة وتحول إلى نسب ذكاء انحرافية متوسطها ١٠٠ وانحرافها المعيارى ١٥ سواء للمقياس الأصلى باقتراح بعض الطرق الإحصائية لحساب الأعمار العقلية المقابلة لدرجات مقياس الأطفال، ورغم أن معيار العمر العقلى ليس من خصائص نسب الذكاء الإنحرافية التى تتألف منها معايير المقياس إلا أنه كان فى ذلك يواجه مطلبا عمليا ملحا نتيجة لشيوع مفهوم العمر العقلى بين خاصة الناس وعامتهم بالنسبة لميدان قياس الذكاء. وقد أعدت بالفعل جداول لتحويل الدرجات الخام فى المقياس إلى أعمار عقلية، إلا أن هذه الحاجة لم تعد ضرورية فى الوقت الحاضر ولذلك نجده فى المقياس المعدل (WISC-R) يستبعد هذه الضرورة، وبالإضافة إلى ذلك فإن المقياس المعدل يتخفف كثيرا من الأسئلة التى لا تتناسب مع الأطفال وأحل محلها أسئلة مشتقة من خبرات الطفولة.

كما زيد طول بعض الاختبارات الفرعية لزيادة الثبات، كما طرأت تحسينات على طرق تطبيق المقياس وإجراءات تصحيحه، وقد أدى ذلك إلى إعطاء متوسطات نسب ذكاء تقترب كثيرا من القيم التى نحصل عليها من مقياس ستانفورد - بينيه (معايير ١٩٧٢)، وبصفة عامة لوحظ أن معايير المقياس المعدل أعلى من معايير المقياس الأصلى.

وتتألف عينة تقنين مقياس وكسلر للأطفال (WISC-R) من ١٠٠ طفل و١٠٠ طفلة فى الأعمار من ٦,٥ سنوات إلى ١٦,٥ سنة، وكان المجموع الكلى للعينة ٢٢٠٠ مفحوصًا صنفت على أساس الإحصاء العام للسكان فى الولايات المتحدة الذى أجرى عام ١٩٧٠ وفى ضوء نفس المتغيرات التى ذكرناها بالنسبة لمقياس الراشدين، وكانت جميعا من الأطفال الأسوياء.

ثبات المقياس: حسبت معاملات الثبات بطريقة التجزئة النصفية لكل اختبار فرعى فى مقياس وكسلر للأطفال ولكل من المقياس اللفظى أو المقياس العملى ، والمقياس الكلى، وقد حسب الثبات بالتجزئة النصفية لكل مجموعة عمرية من المجموعات الإحدى عشرة إلا أن الثبات بطريقة إعادة الاختبار حسب لثلاث مجموعات فقط هى: (أ) من ٦,٥ – ٧,٥ (ب) من ١٠,٥ – ١١,٥، (جـ) من ١٤,٥ – ١٥,٥ وقد بلغ متوسط معاملات الثبات بالتجزئة النصفية ٠,٩٤ للمقياس اللفظى، ٠,٩٠ للمقياس العملى ٠,٩٦ للمقياس الكلى بينما بلغت هذه المعاملات بطريقة إعادة الاختبار ٠,٩٣، ٠,٩٠، ٠,٩٥ على التوالى.

أما معاملات ثبات الاختبارات الفرعية فكانت فى الطبعة المعدلة (طبعة ١٩٧٤) أعلى منها فى الطبعات السابقة، فتراوحت معاملات الثبات لمختلف الأعمار بالتجزئة النصفية بين ٠,٧٠ و ٠,٨٦ وبطريقة إعادة الاختبار بين ٠,٦٥، ٠,٨٨.

وقد تضمن دليل الاختبار برهان على ثبات المقياس فى الطبعة المعدلة ١٩٩١ WISC-III (والذى يصلح للأطفال من سن ٦ إلى سن ١٦ واحدى عشر شهرا) بيانات عن ثبات الاختبارات الفرعية وأيضا لمعاملات ثبات المصححين والذين دربوا إلى درجة كبيرة على كيفية التصحيح.

صدق المقياس: لا تتضمن كراسات تعليمات مقياس وكسلر للأطفال بيانات عن صدقه، ولو أن بعض الشواهد تدل على ذلك فجداول المعايير لكل اختبار فرعى تظهر تمايزا فى العمر دون أن تقدم البيانات للقارئ فى ضوء هذا المحك، وقد قام عدد من الباحثين بحساب الصدق التلازمى للمقياس باستخدام محك التحصيل المدرسى وتراوحت معاملات الارتباط بين ٠,٥٠، ٠,٦٠ وكما هو متوقع كانت معاملات المقياس اللفظى أعلى، كما حسبت معاملات ارتباط مقياس ستانفورد ـ بينيه فبلغت

٠,٧٣ بالنسبة للمقياس الكلى، وكانت معاملات المقياس اللفظى مرة أخرى أعلى.

وتقدم كراسة تعليمات المقياس المعدل معلومات إضافية عن معاملات الارتباط بين الاختبارات الفرعية بعضها وبعض من ناحية، والمقياس اللفظى والمقياس العملى والمقياس الكلى من ناحية أخرى، وقد حسبت هذه المعاملات منفصلة لعينة قوامها ٢٠٠ مفحوص فى كل مجموعة عمرية من المجموعات الإحدى عشرة فى عينة التقنين، وقد بلغ متوسط معاملات ارتباط المقياس اللفظى والعملى ٠,٦٧ ومعنى هذا أن جزئى المقياس بينها قدر مناسب من الاشتراك إلا أن انخفاض معامل ارتباطهما نسبياً يبرر الإبقاء عليهما كمقياسين منفصلين.

أما بحوث التحليل العاملى التى أجريت على المقياس بهدف تحقيق صدق التكوين الفرضى فقد استخرجت عوامل تشبه ما وجد فى تحليل مقياس الراشدين وهى: العامل العام، والفهم اللفظى، والعامل المكانى الإدراكى وعامل الذاكرة، وتأكد أن عامل الفهم اللفظى يرتبط ارتباطا جوهريا بالاختبارات الفرعية الخمسة فى المقياس اللفظى، وأن العامل المكانى الإدراكى يرتبط ارتباطا دالا باختبارى رسم المكعبات وتجميع الأشياء وهما من الاختبارات الفرعية فى المقياس الأدائى العملى.

تقويم المقياس: يمكن القول أن مقياس وكسلر للأطفال مقياس جيد لذكاء الأطفال ويتفق فى الخصائص مع كثير من الاختبارات الفردية، وكانت أنستازى ترى فى طبعة ١٩٦٨ من كتابها (٦١) أن الأساس المنطقى لهذا المقياس ظاهر التناقض، لقد ظهر مقياس وكسلر بلفيو الأصلى عام ١٩٢٩ لمقابلة الحاجة إلى اختبارللراشدين لايوفرها مقياس ستانفورد- بينيه أو أى مقياس آخر يعتمد على مجرد (رفع مستوى) مقاييس الأطفال، وبعدما حقق وكسلر هذا الهدف سعى إلى إعداد مقاييس للأطفال تعتمد فى جوهرها على (خفض مستوى مقاييس الراشدين) بل أنه سار بالتناقض إلى نهايته حينما أعد مقياسا لأطفال ما قبل المدرسة، إلا أن هذا النقد لم يعد له ما يبرره بعد أن ظهرت الطبعة المعدلة من المقياس عام ١٩٧٤ وكذلك تعديل ١٩٩١ وتضمنت محتوى «طفولى» تماما- إلا أن أهم ما يلفت النظر فى هذا المقياس أنه لم يحظ بدراسات تستحق الذكر عنه باللغة العربية ولاتكاد تتوافر بيانات عن معالمه السيكومترية يمكن الاعتماد عليها.

مقياس وكسلر لذكاء أطفال ما قبل المدرسة :

ظهر مقياس وكسلر لذكاء أطفال ما قبل المدرسة (WPPSI) أو Wechsler Preschool and Primary Scale of Intelligence عام ١٩٦٧ ، ويهدف إلى قياس ذكاء الأطفال من سن ٤ سنوات إلى ٦,٥ سنة ، ويتألف من ١١ اختبارا فرعيا يستخدم منها ١٠ اختبارات فقط فى تحديد نسبة الذكاء، ومنها نجد ٨ اختبارات عبارة عن خفض لمستوى الاختبارات المناظرة لها فى مقياس الأطفال WISC آنف الذكر، أما الاختبارات الثلاث الباقية فقد أعدت خصيصا لهذا المقياس، وتصنف الاختبارات الفرعية فى هذا المقياس ـ كما هو الحال فى مقياس WISC, WAIS إلى قسمين: لفظى وعملى، على النحو الآتى (وقد وضعنا علامة * أمام الاختبارات الجديدة) .

المقياس اللفظى	المقياس الأدائى (العملى)
المعلومات	* بيت الحيوان
المفردات	تكملة الصور
الحساب	المتاهات
المتشابهات	*الرسم الهندسى
الفهم	رسم المكعبات
* الجمل (اختبار إضافى)	

ويوصى وكسلر باستخدام الاختبارات اللفظية والعملية بالتناوب لتوفير التنوع والاحتفاظ بميل الطفل وتعاونه، ويتطلب المقياس فى تطبيقه ما بين ٥٠ دقيقة و٧٥ دقيقة فى جلسة واحدة أو جلستين اختباريتين.

وقد حل اختبار الجمل فى هذا المقياس محل اختبار مدى (إعادة) الأرقام فى مقياس WISC ، وفيه يكرر الطفل كل جملة مباشرة بعد عرضها عليه شفويا، وهذا الاختبار يمكن أن يستخدم كبديل لأي اختبار لفظى آخر، أو يطبق كاختبار إضافى لتزويد الفاحص بمعلومات أكثر عن الطفل، وفى كلتا الحالتين لايستخدم فى حساب نسبة الذكاء، أما اختبار بيت الحيوان فيشبه اختبارر موز الأرقام فى مقياس WAIS واختبار الترميز فى مقياس WISC ، ويتضمن الاختبار مفتاحا يعطى للطفل على لوحة يتألف من صور لكلب ودجاجة وسمكة وقطة، لكل منها بيت (اسطوانى ملون

بلون مختلف موضوع تحت الصورة، وعلى الطفل أن يضع الاسطوانة (البيت) فى ذات اللون المناسب فى فتحة تحت صورة الحيوان على اللوحة، وتتحدد درجة الطفل فى ضوء الزمن وعدد الأخطاء وعدد الاستجابات المحذوفة، أما اختبار «الرسم الهندسى» فيتطلب من الطفل محاكاة ١٠ رسوم بسيطة بقلم ملون.

تقنين المقياس: تم تقنين هذا المقياس على عينة تتألف من١٢٠٠ طفلا مصنفة إلى ٦ فئات، مدى كل منها نصف سنة فى الأعمار من ٤ إلى ٦,٥ سنة وكان كل فئة تتكون من ١٠٠ طفل و١٠٠ طفلة اختيرت على أساس بيانات التعداد العام للسكان فى الولايات المتحدة الأمريكية عام ١٩٦٠. وفى تصحيح المقياس يتم تحويل الدرجات الخام إلى درجات معيارية معدلة متوسطها ١٠ وانحرافها المعيارى ٣، أما مجموع الدرجات الموزونة فى المقياس اللفظى والمقياس العملى والمقياس الكلى فيتحول إلى نسبة ذكاء انحرافية متوسطها ١٠٠ وانحرافها المعيارى ١٥. بالإضافة إلى الأعمار العقلية بالمعنى التقليدى.

وقد حسب ثبات الاختبارات الفرعية - فيما عدا اختبار بيت الحيوان - بطريقة التجزئة النصفية، أما اختبار بيت الحيوان فقد حسب ثباته بطريقة إعادة الاختبار لاعتماده الكبير على السرعة فى الأداء، وتراوحت معاملات الثبات للاختبارات الفرعية والمقاييس الكلية بين ٠,٨٠، ٠,٩٢.

أما من حيث صدق المقياس فلا زال الميدان يحتاج لمزيد من البحث، وقد يرجع ذلك إلى حداثة المقياس، وقد حسبت علاقته بمقياس ستانفورد - بينيه فبلغ معامل الارتباط ٠,٧٥ فى مجموعة ٩٨ طفلا تتراوح أعمارهم بين ٥، ٦ سنوات. كما هوالحال فى مقياس WISC فإن معامل ارتباط المقياس اللفظى بمقياس ستانفورد - بينيه أعلى من ارتباط المقياس العملى به. وقد أجريت عدة تحليلات عاملية للمقياس سجلها ساتلر (١٠٢) وفيها تأكد وجود عامل عام بالإضافة إلى عامل لفظى وعامل عملى وكان العاملان الطائفيان أقل ظهورا فى الأعمار الدنيا، وعموما فإن إجراءات تقنين هذا المقياس وحساب ثباته على درجة عالية من الدقة الفنية - كما تقول انستازى (٦١ : ٢٩١) – وهو فى ذلك يتفوق على كثير من اختبارات ذكاء أطفال ما قبل المدرسة التى ظهرت قبل ذلك، ويعكس التطورات التى طرأت على حركة بناء

الاختبارات فى السنوات الأخيرة، ومع ذلك فإن هذا المقياس لم يظهر بأية صورة فى أى بيئة عربية على عكس مقياسى وكسلر السابقين للراشدين والأطفال.

الصور المختصرة من مقاييس وكسلر للذكاء

منذ أصدرت الطبعة الأصلية لمقياس وكسلر – بلڤيو ظهر عدد كبير من المقاييس المعدلة والصور المختصرة لهذا المقياس. وكان الهدف من هذه المقاييس المختصرة هو اختصار للوقت اللازم لتطبيق المقياس مع الاحتفاظ بالحصول على قيمة نسبة الذكاء التى يتم الحصول عليها عند تطبيق المقياس الأصلى وكما تم تحديدها فى المعايير التى تم نشرها فى دليل الإختبارات. وقد تم الإعتماد على حذف بعض الإختبارات الفرعية مع تعديل نسب الدرجات، كذلك تم الوصول إلى اعداد صور مختصرة لحذف بعض المفردات من الإختبارات الفرعية.

وقد قوبلت هذه الصور المختصرة بالقبول وشجع استخدامها فى المواقف التى تتطلب قرارات سريعة بعد التأكد من أن قيم معاملات الإرتباط بين الأداء فى المقياس الأصلى والمقاييس المختصرة بلغ ٠,٩٠ وقد أجريت العديد من الدراسات على اختيار مجموعات الإختبارات الفرعية التى تتكون من اختبارين أو ثلاثة أو أربعة أو خمسة للوصول إلى أفضل مجموعة يمكن أن تحل محل الصورة الكاملة، وبعض من هذه المجموعات تم التحقق من صدقها وثباتها، إلا أن هذه الدراسات لم تجر على عينات ذات طبيعة خاصة مثل المرضى النفسيين أو ضعاف العقول.

وبالرغم من الجهد الكبير الذى بذل فى محاولة الوصول إلى صورة مختصرة، إلا أن هناك العديد من التساؤلات حول الإفتراض الأصلى لإستخدام الصور الأصلية والدرجة الكلية المشتقة منها. هذا بالإضافة إلى أن الملاحظات النوعية التى يمكن الوصول لها عند تطبيق الصورة الكلية لايمكن الوصول لها من الصورة المختصرة.

المقياس البريطانى للقدرة:

تولى مسئولية إعداد هذا المقياس ثلاثة من علماء النفس البريطانيين هم اليوت C.D. Elliott وموراى D.J. Murray وبيرسون L.S. Pearson وظهر باسم المقياس البريطانى للقدرة British Ability Scale وقامت بنشره المؤسسة القومية للبحوث التربوية (NFER).

ويتألف هذا المقياس من ٢٤ مهمة يجد فيها معظم المتخصصين فائدة كبرى فى التشخيص لمستوى واسع من الأعمار ابتداء من سن ٢,٥ سنة وحتى ١٧ سنة.

وتصنف المهام تبعا لمدى الأعمار، فالمهام الخاصة بالمزاوجة الإدراكية per-ceptual matching تقتصر على المدى العمرى من ٢,٥ إلى ٩ سنوات بينما لا تقاس «السرعة فى تجهيز المعلومات» قبل سن ٨ سنوات، ومن أمثلة المهام التى تقيس القدرة الأخيرة أن يعرض على المفحوص خمسة أعداد يتألف كل منها من ثلاثة أرقام ويطلب منه أن يختار أكبرها، وتوجد مهام أخرى لقياس الذاكرة والاستدلال الصورى بمعناه عند جان بياجيه بالإضافة إلى الاستدلال الاجتماعى بالإضافة إلى المعرفة الأساسية بالقراءة والكتابة والسهولة فى التعامل مع الأعداد.

وتتلخص أهم أهداف انشاء هذه المقاييس فيما يلى:

١- التعرف على الأطفال الذى لديهم صعوبات فى التعلم لإمكان توجيههم إلى مدارس تعمل على تنمية قدراتهم.

٢- تحديد صعوبات التعلم وطبيعتها والطرق التى يمكن استخدامها للتغلب عليها.

٣- الإستفادة من هذه المقاييس فى أغراض التشخيص والإرشاد والتوجيه النفسى.

٤- الإستفادة من نتائج تطبيقها فى تحليل ووضع المناهج الدراسية.

٥- اعداد البرامج وطرق التدريس المناسبة للأعمار الزمنية للأطفال.

٦- استخدام المقاييس فى تقدير النمو أو التغير عبر فترة زمنية معينة.

والغرض الرئيسى من هذا الاختبار هو رسم بروفيل للمفحوص، ويمكن للفاحصين استخدام بعض الاختبارات الفرعية دون البعض الآخر، ويرى مؤلفوا الاختبار أن بعض الاختبارات الفرعية يمكن الحصول على متوسط درجاتها وتقدير «نسبة ذكاء عامة» منها والبعض الآخر يعطى «نسبة ذكاء بصرية» والبعض الآخر يعطى «نسبة ذكاء لغوية» وهو فى هذا أشبه بمقياس وكسلر. وفى الوقت نفسه فيه بعض خصائص ستانفورد -بينيه ولعل من أهم السمات المهمة يمكن أن تختار فى ضوء أداء المفحوص ذاته، ومن ذلك مثلا أن اختبار «تصميم المكعبات» فيه يتألف من ١٦ نمطا، ويمكن للطفل الصغير أن يعطى الأنماط الأحد عشر الأولى بينما يبدأ

المراهق من النمط السادس حتى النمط الأخير، ويقدر مستوى قدرة المفحوص من خلال سجل تراكمى لنجاحه وفشله بطريقة راش Rasch ، ويبدو لنا أن هذا المقياس الجديد لايزال فى حاجة إلى مزيد من البحث، كما ندعو الباحثين المصريين إلى محاولة تطويعه للثقافة المصرية خاصة والعربية عامة. مع ضرورة الإشارة أن هذا المقياس اعتمد فى بناءه على عدة نظريات مثل نظرية بياچيه وكولبرج واستخدم نموذج راش فى تدريج المقاييس

مقاييس القدرات الفارقة Differential Ability Scales (DAS)

تعد هذه المقاييس تعديل وامتداد لبطارية المقياس البريطانى للقدرة (BAS) British ABility Scale والتى طورها اليوت C.D. Elliott ، وموراى D.J. Mur-ray، وبيرسون L. S. Pearson سنة ١٩٧٩ . وتشترك هذه البطارية مع مقياس ستانفورد – بينيه ومقياس وكسلر فى أهدافهم لتصنيف الأفراد فى ضوء مستواهم العام فى القدرات، وتحديد البروفيلات لهؤلاء الاشخاص والتى توضح مظاهر القوة والضعف. وتتميز هذه البطارية (DAS) بخصائصها السيكومترية والتطبيقية والتى لا تتوافر فى المقاييس الأخرى، ويؤكد معدو هذا المقياس أنهم لا يهتمون بالذكاء إو نسبة الذكاء ولكن اهتمامهم ينصب على ما يستطيع المفحوص القيام به بالفعل.

ويهدف (DAS) إلى قياس القدرات النوعية الخاصة بدرجة كبيرة من الثبات، كما يعد أداة جيدة فى التشخيص الفارق للفرد والتخطيط للعلاج، وقد تم اختيار مكونات البطارية معتمدا على اطار نظرى مرجعى جيد وكذلك دراسات ميدانية مرجعية، فالأساس النظرى مرن ويعتمد على منظومة هرمية للقدرات العقلية والتى تؤكد عمليا على نمو القدرات المعرفية للمفحوص.

وتتكون بطارية (DAS) من ٢٠ اختبارا فرعيا مصنفة إلى ثلاثة مكونات رئيسية هى: أ: الإختبارات المحورية (Core tests) وعددها ١٢ اختبارا.

ب: الاختبارات التشخيصية وعددها خمسة اختبارات.

جـ: الإختبارات التحصيلية وعددها ثلاثة اختبارات.

وتشكل الإختبارات فى الفئة الأولى والثانية الإختبارات المعرفية. والإختبارات

المحورية تتشبع تشبعا عاليا على العامل العام، بينما الاختبارات التشخيصية منفردة لا تتشبع على العامل العام مما يؤكد قياسها لقدرات خاصة لا تتجمع فى عوامل. ويوضح جدول (٢٦) هذه المجموعات.

جدول (٢٦)

تصنيف مقاييس القدرات الفارقة DAS

الإختبارات المحورية	الإختبارات التشخيصية	الإختبارات التحصيلية
١- بناء المكعبات ٢- الفهم اللفظى ٣- الصور المتشابهة ٤- تسمية الكلمات ٥- المفاهيم العددية الأولية ٦- نسخ النموذج ٧- بناء النماذج ٨- استدعاء الأشكال ٩- تعريف الكلمات ١٠- المصفوفات ١١- المتشابهات ١٢- المتتاليات الاستدلال الكمى	١- مزاوجة الحروف/ الاشكال ٢- تذكر الاعداد ٣- تذكر الأشياء ٤- التعرف على الصور ٥- سرعة تجهيز المعلومات	١- المهارات العددية الأساسية ٢- التهجى ٣- قراءة الكلمات

وتطبق هذه الاختبارات طبقا لنظام محدد ولمدى عمرى محدد أيضا. ويمكن الرجوع إلى دليل الإختبار ذاته لمعرفة الحدود الزمنية لكل اختبار. كما توجد بعض الاختبارات خارج المدى الزمنى للإختبار يمكن تطبيقها على المفحوصين الفائقين أو من الضعفاء ويلاحظ أن الإختبارات التحصيلية يبدأ تطبيقها من سن خمس سنوات ويستمر استخدامها فى مدى عمرى كبير (حتى سن ١٧ سنة) ويوضح الشكل التالى مستوى الإختبارات والعمر الزمنى وامتداده فى الأعمار المختلفة.

بطارية الاختبارات المعرفية	الأعمار الزمنية	مرحلة ما قبل المدرسة										مراحل التعليم				
		سنة	شهر	سنة	شهر	سنة	شهر	سنة	شهر	سنة	شهر	6	7	8	9	الى17
	من	2	6	3	-	3	6	4	•	4	6	5				
	الى	2	11	3	5	3	11	4	5	4	11					
	الاختبارات المحورية															
	بناء المكعبات	x	x	/		/		/								
	الفهم اللفظى	x	x	x		x		x		x		••				
	الصور المتشابهة	x	x	x		x		x		x		••	••			
	تسمية الكلمات	x	x	x		x		x		x		/	/	••		
	المفاهيم العددية الأولية	*	*	x		x		x		x		/	••			
	نسخ النموذج			x		x		x		x		/	/			
	بناء النموذج		*	x		x		x		x		x	x	x	x	x
	استدعاء الأشكال									/		x	x	x	x	x
	تعريف الكلمات									*		x	x	x	x	x
	المصفوفات									/		x	x	x	x	x
	المتشابهات									*		x	x	x	x	x
	المتتاليات والاستدلال الكمى									*		x	x	x	x	x
	اختبارات التشخيص															
	مزاوجة الحروف والأشكال							*		x		••	••			
	تذكر الأعداد	*	x	x		x		x		x		x	x	x	x	x
	تذكر الأشياء							x		x		x	x	x	x	x
	التعرف على الصور	*	x	x		x		x		x		x	x	••	••	••
	سرعة تجهيز المعلومات									*		x	x	x	x	x

	اختبارات التحصيل												
	المهارات العددية									x	x	x	x
	التهجى									x	x	x	x
	قراءة الكلمات							*		x	x	x	x

الرموز المستخدمة الدلالة على صلاحية الاختبارات لأعمار معينة	
مستوى العمر العادى	x
مستوى العمر ممتد	/
أعلى من مستوى العمر	*
أدنى من مستوى العمر	••

شكل (3):بنية اختبارات القدرات الفارقة

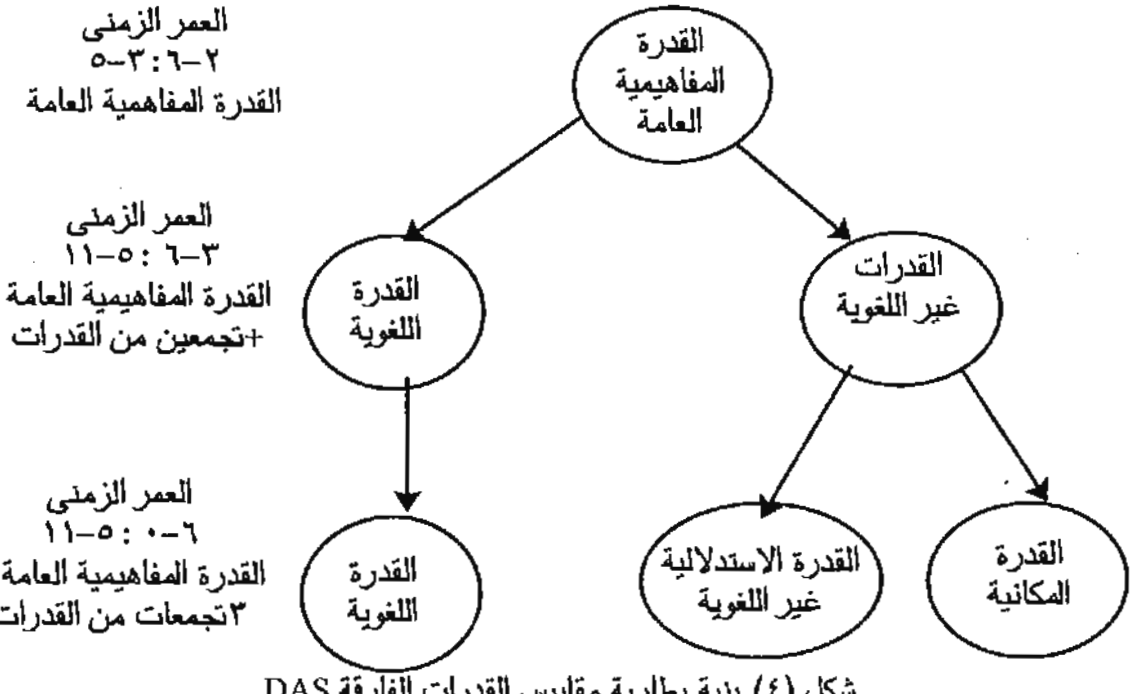
عن أنستازى وأوربانا (1997-228)

وبصفة عامة يمكن القول إن هذه البطارية تم اعدادها فى ضوء ما توصلت له بطارية BAS كأداة جديدة بمعالم حديثة فى القياس والتى تعكس خبرة ٢٠ عاما فى تطوير المقياس (من عام ١٩٦٠-١٩٨٠).

وقد تم تقنين DAS على عينات قوامها ٣٤٧٥ مفحوصا مع الإحاطة أنه اختبار فردى لمفحوصين تمتد أعمارهم من سن ٢-٦ إلى ١٧-١١ سنة واستمر التطبيق وجمع البيانات واختيار العينات التى تمثل الخصائص الديموجرافية المستحبة من التعداد الرسمى للسكان، هذا بالإضافة إلى العينات من ذوى الإحتياجات الخاصة. وقد اعتمد فى انتقاء المفردات نظرية الاستجابة للمفردة الاختبارية One- parameter model of item Response Theory (IRT) ، وعلى ذلك يمكن للمفحوص أن يختبر بالمفردات المناسبة لقدرته (Adeptive Testing) وتحسب درجته بناء على عدد المفردات وصعوبتها التى أجاب عليها.

وقد تحدد ثبات البطارية باستخدام طريقة إعادة التطبيق بفاصل زمنى مدته أسبوعين وحتى سبع أسابيع وتراوح معامل الإرتباط بين ٠,٧٩ ، ٠,٩٤ للإختبارات الفرعية و٠,٦٠ ، ٠,٩٠ ، كما تم حساب معامل الإتساق الداخلى باستخدام نموذج الإستجابة المفردة IRT.

أما صدق البطارية فقد تم تحديده باستخدام التحليل العاملى التوكيدى للوصول إلى بنية البطارية، وتبين أن بنية الاختبار تختلف بزيادة العمر وأن القدرات تصبح فارقة، كما تم أيضا حساب معاملات الإرتباط بين الأداء فى هذه البطارية والأداء فى اختبارات بينية ووكسلر. كما أثبتت الدراسات على هذه البطارية أنها هرمية البنية. ويوضح الشكل التالى هذه البنية فى الأعمار المختلفة.



شكل (٤) بنية بطارية مقاييس القدرات الفارقة DAS

منظومة التقييم المعرفى م.ث.م (الكاس) المرتكزة على نموذج PASS

وقد ظهرت أداة جديدة لتقويم الأداء الفردى للعمليات المعرفية فى نهاية التسعينيات من القرن الماضى وهى (CAS) Das - Naglieri Cognitive Assessment system. وقد طور هذه الأداء فى أكثر من عشر سنوات، وقد بنيت وتطورت من خلال التكامل النظرى والتطبيقى فى مجال علم النفس المعرفى وتستخدم لقياس التخطيط – الإنتباه – التآنى – التتابع Planing - Attention - Simultaneity - Succession ، وقد اعتمد بناء هذا المقياس على نموذج الذكاء الذى اقترحه كل من Das J. P & Jack A. Naglieri معتمدين فى ذلك على الأساس النظرى المعرفى والتنظيم العقلى عند لوريا A. R. Luria (عن انستازى وأوربانا ١٩٩٧: ٢٣٣).

وقد حددت مهام القياس لقياس الوظائف المعرفية المتضمنة فى عملية التعلم وغير المعتمدة على المدرسة والتى تتضمن المهام الأربعة المشار لها (التخطيط – الانتباه – التآنى – والتتابع) وتتضمن المهام أنشطة لغوية وأخرى غير لغوية وتقدم من خلال مواد بصرية وسمعية.

ويهدف المقياس إلى «تقدير» Assess الاستراتيجيات المتضمنة فى تحقيق المهام المتضمنة فيه ، والمقياس يصلح للأعمار من ٥ سنوات وحتى ١٧ سنة واحدى عشر شهرا، وصمم خصيصا لربط تقدير الأداء بالتدخل كالآتى:

١– تقدير أداء الأفراد الذين يعانون من صعوبات فى التعلم.

٢– تقدير أداء الأفراد الذين يعانون من اعاقات ذهنية.

٣– تقدير أداء الأفراد الذين يعانون من اضطرابات فى الانتباه واضطراب النشاط الزائد.

٤– تقدير أداء الأفراد الذين يعانون من اضطرابات انفعالية شديدة.

٥– تقدير أداء الأفراد الذين يعانون من مشكلات فى التخطيط.

٦– تقدير أداء الأفراد الذين لديهم اصابات أو تلف فى المخ.

٧– تقدير أداء الموهوبين.

وتتكون البطارية من اختبارات اساسية وأخرى فرعية، وتقدر الدرجات فى الاختبارات الفرعية وتحول إلى درجة معيارية متوسطها ١٠ وانحراف معيارى ٣، أما البطارية القياسية فإنها تعطى درجة متوسطها ١٠٠ وانحرافها المعيارى ١٥ .

وقد خضعت منظومة التقييم المعرفى Casإلى تقنين جيد نظرا لما تتمتع به من اعداد جيد واطار نظرى ونشر عن ذلك تقارير متعددة، مما يؤكد أهمية هذه البطارية فى تقدير الأداء المصرفى.

وقد قام أيمن الديب محمد شوشة تحت اشراف صفاء الأعسر فى مصر باقتباس واعداد هذه البطارية ونشرت عام ٢٠٠٦ فى اخراج جيد، كما أجريت بحوث للماجستير والدكتوراه * واستخدمت فيها هذه البطارية، ويتطلب الأمر اجراء البحوث التى تستخدم هذه البطارية لضمان فعالية استخدامها فى المجتمع المصرى والعربى.

بطارية كوفمان:

فى عام ١٩٨٣ أصدر كوفمان وزوجته اختبارا مهما باسم بطارية الأطفال لكوفمان Kaufman Assessment Battery for chidren بحيث تصلح من سن ٢,٥ إلى ١٢ سنة وتتألف البطارية من ١٦ اختبارا، إلا أنها مثل المقياس البريطانى للقدرة، تخصص بعض الاختبارات الفرعية للأطفال الصغار والبعض الآخر للأطفال الكبار.

وتقوم الأسس للبطارية على الاتجاهات الحديثة فى مجال سيكولوجية الجهاز العصبى وعلم النفس المعرفى، ولهذا يمكن الحصول منه على أربع درجات «كلية» رئيسية هى:

١ – درجة التجهيز التتابعى Sequential Processing ويمكن الحصول عليها من أداء المفحوص لمهام الذاكرة والتى تتطلب منه إعادة إنتاج سلسلة من الأعداد بعد الاستماع إليها، أو إصدار سلسلة من حركات اليد بعد رؤيتها أو لمس رسوم يسهل التعرف عليها بنفس الترتيب الذى يستخدمه الفاحص فى تسمية الأشياء المصورة فى هذه الرسوم.

٢ – درجة التجهيز المتآنى Simultaneous Processing ويمكن الحصول عليها من أداء مهام تشبه ما تتضمنه اختبارات مشابهة لترتيب الصور فى مقياس وكسلر،

* : راجع قائمة المصادر.

ولاختبار المصفوفات، واختبار الإغلاق (الجشطالتى)، ومن الاختبارات الجديدة المستحدثة فى هذه البطارية اختبار يسمى «النافذة السحرية» وفيه يحرك الفاحص صورة بحيث تمر فى نافذة ضيقة ويطلب من المفحوص أن يتعرف على الشىء المصور.

٣ – الدرجة المركبة للتجهيز العقلى وتمثل متوسط الدرجتين السابقتين، وتمثل الدرجة التى يحصل عليها المفحوص فى الاختبارات الفرعية التى يتم تطبيقها عن طريق التمثيل الصامت ما يسمى «الدرجة غير اللفظية».

٤ – درجة التحصيل ويمكن الحصول عليها من أداء المفحوص على المهام التى تعتمد مباشرة على الخبرة الناتجة عن الثقافة أوالمدرسة، ويشمل ذلك القراءة والحساب وتسميه الصور التى تدل على أشخاص أو أماكن معروفة (اختبار المعلومات) أوتدل على أشياء (اختبار المفردات).

وينعكس فى هذه البطارية التمييز الشهير الذى اقترحه ريموند كاتل بين الذكاء «السائل» والذكاء «المتبلور» وخاصة فى درجتى التجهيز المتآنى والتحصيل على التوالى، أما التجهيز التتابعى فيدل على الكفاءة فى الانتباه والتسميع (البروفة الذهنية) وهى عمليات مستقلة عن الاستدلال والمعلومات ، ولايركز كوفمان على الدرجات المنفصلة للاختبارات الفرعية الفردية (كما هو الحال فى المقياس البريطانى للقدرة)، ولكنه يقترح تفسير أى اختلاف ملحوظ داخل كل فئة من الفئات العامة.

ويستخدم المقياس فى معايير الاختبارات الفرعية للتجهيز درجة معيارية معدلة متوسطها ١٠ وانحرافها المعيارى ٣ كما هو الحال فى الاختبارات الفرعية لوكسلر، أما بالنسبة للدرجات الكلية من ناحية ودرجة الاختبارات الفرعية للتحصيل من ناحية أخرى فإن متوسطها ١٠٠ وانحرافها المعيارى ١٥. ويمكن للدرجات أن تتحول أيضاً إلى مئينيات، والتى حسبت على المستوى القومى من ناحية وعلى أساس المستويات الاقتصادية الاجتماعية من ناحية أخرى.

ويمكن للطفل أن يستجيب على معظم اختبارات هذه البطارية بالإشارة أو غيرها من الاستجابات غير اللفظية، ولاتوجد مهمة فى البطارية كلها يتطلب أكثر من كلمتين كاستجابة لفظية، ولهذا يقترح المؤلفان استخدام اختبار الفهم من مقياس وكسلر كمكمل لاختبارهما، كما يؤكد دليل البطارية على أن الاختبارات تقيس «النشاط العقلى الراهن» وليس القدرات الفطرية.

تقنين مقياس كوفمان لذكاء الأطفال:

قنن اختبار كوفمان للأطفال على مستوى قومى، وبلغ عدد عينات التقنين ٢٠٠٠ مفحوصا يتراوح عمرهم بين سنتين وستة شهور – اثنى عشر سنة وستة شهور بالإضافة إلى مجموعات من السود والبيض من مستويات اقتصادية اجتماعية مختلفة للوصول إلى معايير مرتبطة بالسلالات (Race) المختلفة، وأيضا مستويات مختلفة من تعليم الأباء. وقد استخدمت المعايير المشتقة من أداء عينات التقنين كوسيلة مساعدة فى التشخيص والتفسير، كما تم الاستفادة من التقنين على عينات من ذوى الاحتياجات الخاصة، والأطفال غير الناطقين باللغة الانجليزية (كوفمان وكوفمان ١٩٩٠ فى : أناستازى وأوربانا (٦١: ٢٢٣).

مقياس كوفمان للمراهقين والكبار KAIT

تم انشاء البطارية لقياس الذكاء فى الأعمار من سن ١١–٨٥ سنة، وتتضمن البطارية اختبارات تعتمد على الذكاء السائل وأخرى على الذكاء المتبلور كما حدده هورن Horn وكاتل Cattell (١٩٩٦) بإنطباعات عن ذكاء الكبار كما اقترحته نظريات أخرى (بياجيه ١٩٧٢، لوريا ١٩٨٠، جولدن ١٩٨١ فى اناستازى وأوربينا ٦١ : ٢٢٤).

وتتضمن الاختبارات التى تقيس الذكاء المتبلور المفاهيم المكتسبة من التعليم والثقافات المتعددة، وأيضا الذكاء السائل كما يرتبط بحل المشكلات الجديدة. والبطارية المحورية (Core) تتكون من ثلاثة اختبارات فى كل مقياس، كما تتضمن البطارية اختبارات اضافية للأشخاص الذى يشتبه فى اصابتهم فى الجهاز العصبى يمكن أيضا اضافتها مع أربعة اختبارات أخرى. وتتضمن البطارية أيضا اختبارا يحدد المكانة العقلية لتقدير مدى الإنتباه والتوجه للمفحوصين الذين لا يستطيعون تطبيق البطارية كاملة عليهم.

ويوضح دليل البطارية الخصائص السيكومترية والتى تتصف بها مقاييس الذكاء الجيدة من صدق وثبات كما أنها سهلة التطبيق والتصحيح بالإضافة إلى أن مفردات الإختبارات شيقة المحتوى.

البطارية المختصرة لكوفمان (K- BIT)

أهتم كوفمان بإنشاء بطارية مختصرة Kaufman Brief Intelligence Test يسهل تطبيقها وتصحيحها ويمكن استخدامها فى مدى عمرى واسع (٤-٩٠ عاما) وقد تم تقنين هذه البطارية على ٢٠ ٪ من العينة الأصلية للتقنين (KAIT) والتى تضمنت ٢٠٠٠ مفحوصا.

وتتألف البطارية من ثلاثة اختبارات فرعية، واحد منها لفظى يتضمن كلمات وتعريفات، والإختبار الثانى يتضمن مصفوفات غير لفظية، ويمكن الحصول على نسبة ذكاء انحرافية من الدرجات التى يحصل عليها المفحوص من اختبار الكلمات والمصفوفات والدرجة الكلية لهما.

قياس ذكاء الرضع والأطفال فى سن ما قبل المدرسة

توجد فئة من اختبارات الذكاء الفردية أعدت خصيصًا لقياس ذكاء الرضع والأطفال فى سن ما قبل المدرسة. والواقع أن قياس الذكاء فى هذه المرحلة له خصائصه ومشكلاته، ولعل أهمها تدريب الفاحص تدريباً جيدا على أساليب الملاحظة العلمية المباشرة. وبالطبع يمكن تطبيق بعض الاختبارات الجماعية على أطفال هذه المرحلة، إلا أنه بوجه عام يمكن القول أن الاختبارات الجماعية لا تصلح إلا بعد وصول الطفل إلى سن المدرسة (٦ سنوات). ويترتب على ذلك أن معظم اختبارات الذكاء فى هذا السن (إن لم يكن كلها) من النوع العملى الشفوى وتتناول فى جوهرها النمو الحس حركى، والخلاصة أن اختبار الأطفال فى هذه المرحلة (وبخاصة بعدما يصل الطفل إلى مستوى من النمو اللغوى يسمح بالاتصال والتواصل إنما هو عملية تفاعل بين الأفراد inter personal من نوع رفيع، ونعرض فيها يلى بعض الاختبارات المستخدمة فى هذه المرحلة وقد يكون أشهرها هو مقياس وكسلر لذكاء أطفال ما قبل المدرسة والذى آثرنا عرضه مع باقى البطارية الخاصة بوكسلر. ونعرض فيما يلى بعض هذه الاختبارات.

الجداول الارتقائية لجيزل:

من المحاولات الكثيرة التى بذلت لقياس ذكاء الرضع وأطفال ما قبل المدرسة ما يسمى المقاييس الارتقائية devlopmental scales التى تعتمد على نتائج البحث فى علم النفس الارتقائى، ومن أشهر هذه المقاييس جداول جيزل Gesell schedules التى ظهرت عام ١٩٤٧ وتفيد فى تحديد مستوى النمو الذى يصل إليه الطفل (بالمعايير الأمريكية بالطبع). وتعتمد فى تقدير جوانب النمو إجراء تقنين يعتمد على الملاحظة المباشرة لاستجابات الطفل لمثيرات مقننة، وتقويم سلوكه ونموه فى مواقف الحياة اليومية فالملاحظة جوهر هذه الأداة بالإضافة إلى المعلومات التى يمكن الحصول عليها من مصادر أخرى كالأم، وتشمل هذه الجداول فترة طويلة نسبيا من النمو حيث تمتد من سن ٤ أسابيع حتى سن ٥ سنوات، وتؤكد نتائج البحث أن هذه الجداول تعطى معاملات ثبات بين الفاحصين المدربين يزيد على ٠,٩٥ على الرغم من أن نظام الملاحظة والتصحيح أقل تقنينا إذا قورن بغيرها من الاختبارات فتدريب الملاحظين يرفع الثبات (٦٣: ٢٥٦)، وقد استخدمت هذه الجداول منذ ظهورها من جانب الأخصائيين النفسيين والأطباء العقليين سواء فى البحث العلمى أو الممارسة العملية، وتشمل هذه الجداول فى صورتها الأخيرة خمسة حالات رئيسية هى: الحركات الغليظة، الحركات الدقيقة، اللغة، السلوك التوافقى، والسلوك الشخصى الاجتماعى والتى يزود الفاحص أو الملاحظ مع كل منها بوصف لفظى مفصل للسلوك المعتاد فى المراحل العمرية المختلفة بالإضافة إلى رسوم توضيحية لهذا السلوك.

خفض مستوى مقياس ستانفورد - بينيه:

أدى هذا الاهتمام بذكاء الرضع وأطفال ما قبل المدرسة إلى إدخال تعديلات جوهرية على مقاييس بينيه بحيث يمكن «خفض مستواها»، وكانت أولى هذه المحاولات ما قام به كولمان Kuhlmann عام ١٩٢٢ حينما عدل مقاييس بينيه الأصلية بحيث تصل إلى سن ثلاثة شهور، وكان اختباره من المحاولات المبكرة لقياس ذكاء الرضيع ثم عدل هذا المقياس عام ١٩٣٩ ظهر بإسم مقياس كولمان - بينيه بحيث يشمل معظم فترة الطفولة المبكرة.

من أهم المقاييس التى استخدمت أسلوب خفض مقياس ستانفورد - بينيه مقياس

كاتل الذى يسمى مقياس كاتل لذكاء الرضيع Cattell Infant Intelligence Scale الذى يعتبره كثير من علماء النفس ملائما لقياس ذكاء الرضع.

وقد أعد هذا المقياس بغرض «خفض مستوى» تعديل عام ١٩٣٧ من مقياس ستانفورد - بينيه، وفيه استخدم كاتل Cattell أسئلة ستانفورد - بينيه بالإضافة إلى أسئلة أخرى أنشئت أو اقتبست من جداول جيزل وغيرها من اختبارات الرضع المتاحة، وصنفت الأسئلة إلى مستويات عمرية، وحسب العمر العقلى ونسبة الذكاء على النحو الذى استخدم فى طبعة ١٩٣٧ من مقياس ستانفورد - بينيه، ويشمل هذا المقياس مدى من العمر يمتد بين شهرين إلى سن ٣٠ شهرا، وخلال العام الأول تتحدد مستويات العمر فى فترات طول كل منها شهر واحد، وفى العام الثانى تتحدد بفترات طولها شهران، أما فى النصف الأول من العام الثالث فإن طول الفترات ٣ شهور، فإذا اجتاز الطفل جميع الأسئلة المخصصة لهذه الأشهر الثلاثين الأولى من الحياة يستمر الفاحص معه بتطبيق مقياس ستانفورد - بينيه مبتدئا بمستوى ٣ سنوات.

وقد قنن مقياس كاتل لذكاء الرضع على مجموعة من ٢٧٤ طفلا، أعيد اختبار كثير منهم فى الأعمار ٣، ٦، ٩، ١٢، ١٨، ٢٤، ٣٠، ٣٦ شهرا.

ومن الاختبارات التى تنتمى إلى نفس الفئة اختبارات ميريل - بالمر Marrell Palmer واختبارات العالم الإنجليزى فالنتاين C.W. Valentine والجداول الارتقائية التى أعدها فى بريطانيا العالم الإنجليزى جريفيت Griffith وفى رأى فرنون (١١٤: ٦٣) أن جداول جريفيت هذه هى أكثر اختبارات ذكاء الرضع شمولا وأدقها تقنينا، فهى تتضمن ٣ أسئلة لكل أسبوع من أسابيع العمر فى العام الأول من الحياة، وسؤالين لكل أسبوع فى العام الثانى، وبهذا يبلغ عدد أسئلة المقياس ٢٦٠، وتصنف إلى ٥ مجموعات، كل منها متكافئ مع الفئات الأخرى فى الدرجات بحيث يعطى المقياس تخطيطا نفسيا (بروفيلا) للنمو فى كل فئة، بالإضافة إلى درجة كلية تسمى النسبة العامة General Quotient .

ويتضمن الجدول رقم (٢٦) هذه الفئات الخمسة، مع أمثلة منها فى بداية المقياس ونهايته.

جدول رقم (٢٧) فئات مقياس جريفيت وأمثلة من أسئلته

الفئـــة	أمثلة من سلوك الطفل بعد الولادة	أمثلة من أسئلة المقياس عند اقتراب الطفل من سن عامين
(أ) النشاط الحركى	الرفس. رفع الرأس	النزول على الدرج دون مساعدة أحد
(ب) النشاط الاجتماعى	الابتسام، التعرف على الأم	توجيه أسئلة إليه عن أشياء موضوعة أمامه على منضدة.
(جـ) السمع والكلام	الإجفال من الضوضاء	استخدام جمل تتألف من ٤ كلمات أو أكثر
(د) تآزر العين واليد	تتبع أضواء معينة	قذف كرة فى سلة
(هـ) الأداء العملى	إمساك الأشياء	تجميع أجزاء لعبة

مقاييس بيلى لنمو الرضع:

شهد عقد الستينات والسبعينات اهتماما جديدا بالاختبارات التى تصلح للرضع وأطفال ما قبل المدرسة، ومن أهم العوامل التى ساعدت على ذلك التوسع السريع فى البرامج التعليمية المعدة للأطفال المتخلفين عقليا من ناحية والنمو الهائل فى برامج ما قبل المدرسة فى مجال التربية التعويضية للأطفال المحرومين اجتماعيا وثقافيا وقد ظهرت اختبارات جديدة لمواجهة هذه المطالب الملحة.

ومن أشهر هذه الاختبارات مقاييس بيلى لنمو الرضع Bayley Scales of Infant Development والذى صدر فى عام ١٩٦٩ والذى اعتمد فى اعداده على الدراسات الطوليه، ويشمل مواد من جداول جيزل ومن بعض الاختبارات الأخرى التى تصلح لهذه المرحلة. وتتضمن مقاييس بيلى ثلاث أدوات متكاملة لتقدير المستوى الارتقائى للطفل من سن شهرين إلى ٢,٥ * سنة وهى: المقياس العقلى والمقياس الحركى وسجل سلوك الرضع. ويشمل المقياس العقلى عمليات الإدراك والذاكرة والتعلم

* ظهرت الطبعة الثانية للمقياس Baylay II عام ١٩٩٣ لتقدير المستوي الارتقائي للطفل للفترة من شهر واحد وحتي ثلاث سنوات ونصف.

وحل المشكلات والتلفظ، وبدايات التواصل اللفظى، والتفكير المجرد فى صوره المبكرة. ويشمل المقياس الحركى القدرات الحركية الغليظة ومهارات المعالجة بالأيدى والأصابع، ويتداخل هذا المقياس مع المقياس السابق حيث يلعب النمو الحركى دوراً مهما فى تفاعل الطفل مع البيئة وبالتالى فى نموه العقلى، أما سجل سلوك الرضيع فيستخدمه الفاحص أو الملاحظ لتقويم الجوانب المختلفة فى نمو الشخصية مثل السلوك الانفعالى والاجتماعى، ومدى الانتباه، والمثابرة، والتوجه نحو الهدف.

وقد أعدت لهذه المقاييس معايير مشتقة من أداء عينة ممثلة للمجتمع الأمريكى من ١٢٦٢ طفلا موزعين بالتساوى تقريبا حسب الأعمار من شهرين إلى ٣٠ شهرا واستبعدت من العينة حالات أطفال المؤسسات والأطفال المبتسرين والأطفال الذين يعيشون فى بيئات أسرية مزدوجة اللغة، ويعطى المقياس العقلى والمقياس الحركى مؤشرين منفصلين وتتحول الدرجات الخام فيهما إلى درجات معيارية اعتدالية متوسطها ١٠٠ وانحرافها المعيارى ١٦ (كما هو الحال فى مقياس ستانفورد - بينيه) - وقد أعدت هذه المؤشرات الارتقائية كمعايير داخل المجموعات العمرية مصنفة على فترات طولها نصف شهر خلال الأعمار من شهرين إلى ٦ شهور، وعلى فترات طولها شهر واحد للأعمار من ٦ شهور إلى ٣٠ شهرا.

وحسبت معامل الثبات لهذه المقاييس باستخدام طريقة التجزئة النصفية لكل مجموعة عمرية، وقد بلغ وسيط هذه المعاملات للمقياس العقلى ٠,٨٨ وللمقياس الحركى ٠,٨٤ ويعطى المقياس معامل اتفاق مرتفع بين الملاحظين والفاحصين. أما الصدق فلا تتوافر عنه أدلة كافية، ولو أن بيلى يرى أن استخدام مقاييسه أفضل ما يكون لأغراض التشخيص وليس التنبؤ.

مقاييس مكارثى لقدرات الأطفال:

من المقاييس الحديثة والمهمة أيضًا فى هذه المرحلة مقياس مكارثى لقدرات الأطفال MaCarthy Scales of Children Abilities والذى يصلح للأطفال فى الأعمار من ٢,٥ إلى ٨,٥ سنة، ويتألف من ١٨ اختبارات مصنفة إلى ستة مقاييس متداخلة هى: المقياس اللفظى والإدراكى الأدائى، والكمى، والمعرفى العام، ومقياس الذاكرة، ومقياس الحركة.

ويقترب المقياس المعرفى العام من فكرة المقياس التقليدى الكلى للنمو العقلى، ويتم الحصول عليه من مجموع درجات المقاييس الثلاثة الأولى والتى تؤلف ١٥ اختبار من ١٨ اختبار تؤلف البطارية الكلية، وتتحول الدرجات الخام فى هذا المقياس إلى درجات معيارية متوسطها ١٠٠ وانحرافها المعيارى ١٦ حسبت لكل مجموعة عمرية مداها ٣ شهور، أما درجات المقاييس المنفصلة فتتحول إلى درجات معيارية اعتدالية متوسطها ٥٠ وانحرافها المعيارى ١٠ لنفس المجموعات العمرية.

وقد بلغت عينة تقنين هذا المقياس ١٠٠ طفل تقريبا لكل مجموعة عمرية من المجموعات العشر التى تمتد على فترات مداها نصف سنة للمستويات العمرية من ٢,٥ – ٥,٥ سنة ، ومداها عام كامل للمستويات العمرية من ٥,٥ – ٨,٥ سنة، وبلغ قوام العينة الكلية ١٠٣٢ طفلا سويا، تمثل المجتمع الأمريكى حسب تعداد ١٩٧٠.

وحسب معامل الثبات بطريقة التجزئة النصفية فبلغ فى المتوسط ٠,٩٢ للمجموعات العمرية المختلفة فى المؤشر المعرفى العام، وبالنسبة للمقاييس الخمسة الأخرى يتراوح معامل الثبات بين ٠,٧٩ ، ٠,٨٨ واستخدمت أيضًا طريقة اعادة الاختبار (ن = ١٢٥ طفلا) فبلغ معامل الثبات للمقياس المعرفى العام ٠,٩٠ وتتراوح للمقاييس الخمسة الأخرى بين ٠,٦٩ ، ٠,٨٩.

أما الصدق فلا تتوافر عنه أدلة كافية باستثناء بعض البيانات عن الصدق التنبؤى باستخدام محك التحصيل المدرسى فى نهاية الصف الأول بالمدرسة الابتدائية. وعموما البطارية تعتمد فى بنائها الأصلى على مصادر متعددة: الخبرة الكلينيكية ونتائج البحث فى علم نفس النمو ونتائج بحوث التحليل العاملى.

وقد أجرى العديد من الدراسات الطولية لمقاييس مكارثى فى الفترة من ظهور المقاييس فى السبعينيات استمرت لمدة ٢٠ عاما من البحث وبصفة خاصة الدراسات التى قام بها Valencia (١٩٩٠ فى ٦١) على ثبات وصدق المقاييس. وكذلك الدراسات التى اجريت على ذات المقاييس فى مجال التشخيص الكلينيكى ونشرها كوفمان وكوفمان عام ١٩٧٧ والتى مازال الاهتمام بها كبيرا فى مجال التشخيص كذلك تتمتع مقاييس مكارثى بخصائص سيكومترية دقيقة وبصفة خاصة ما يرتبط بالأطفال فى مرحلة ما قبل المدرسة.

الإتجاه النمائى فى قياس وتفسير ذكاء الأطفال

المقاييس المعتمدة على نموذج جان بياجيه:

تختلف وجهة نظر بياجيه عن الذكاء اختلافا تاما بالمقارنة بأصحاب النماذج العاملية، فقد أهتم بياجيه بالتعرف على الإستجابات الخاطئة التى يجيب بها الأطفال فى اختبارات بينية، ومن خلال ملاحظاته المتكررة لسلوك الأطفال واستجاباتهم فى مواقف التفكير والإستدلال أمكن لبياجيه أن يستخلص أن هناك نظام عقلى يحكم تفكير الأطفال ويختلف عن تفكير الكبار.

وقد بذل بياجيه جهدا كبيرا فى تحديد التراكيب المعرفية التى تميز كل مرحلة عمرية، ويتضمن نموذجه مجموعة من المفاهيم التى تصف وتفسر العمليات العقلية وهى مفهوم الوظائف العقلية والمحتوى المعرفى والبنية المعرفية والعمليات العقلية. وهذه المفاهيم تبدأ بأبسط وحدات السلوك الذى يمكن ملاحظته وتنتهى بالوحدة ذات الإستقلالية النسبية وهى المرحلة أو مفهوم المراحل*.

وقد ظهرت فى السنوات الأخيرة مقاييس لمراحل العمر المختلفة معتمدة على نتائج البحوث الرائدة التى قام بها عالم النفس السويسرى جان بياجيه لفترة تتجاوز نصف قرن، ومن السمات المهمة فى هذه المقاييس أنها من النوع الرتبى ordinal ، وقد عرضنا بعض خصائص هذا النوع من المقاييس فى الفصلين الأول والسادس، وبإيجاز نقول أنها تفترض التتابع المنظم والواحد للنمو خلال المراحل المتتابعة، وهى المراحل التى تمتد من مرحلة الرضاعة حتى مرحلة المراهقة وما بعدها، والتى يسميها بياجيه العيانية، ومرحلة العمليات الشكلية الصورية (راجع تفصيل هذه المراحل فى ١٩٩٩:٣٤ المرجع رقم ٢٩)، وتتسم هذه الاختبارات أيضاً بأنها مرجعة إلى المحتوى وتنسب إلى المستوى فى التفسير وليس إلى المعيار، لأنها تقدم وصفا كيفيا لما يستطيع الطفل أداءه فى كل مرحلة، والمهام التى تتألف منها هذه الاختبارات تهدف إلى التعرف على مفاهيم معينة لدى الطفل، أو ما يسميه بياجيه المخططات المعرفية Cog-nitive Schemata (٢٩) وليس قياس سمات واسعة النطاق، وبالنسبة لتطبيق هذه

---

* لمزيد من التفصيل ارجع الى فؤاد أبو حطب، آمال صادق: نمو الإنسان من مرحلة الجنين إلى مرحلة المسنين سنة ١٩٩٩، نادية شريف: الاتجاهات المختلفة حول تفسير الذكاء ١٩٩٥ .

الاختبارات فإنه يتطلب الحصول على تفسير الطفل لحدث معين ملاحظ وأسباب هذا التفسير ويعتمد التصحيح فى جوهره على نوعية الاستجابات وكيفها وليس على عدد الاستجابات الصحيحة لعدد من الأسئلة، كما هو الحال فى الاختبارات التقليدية، ولهذا نجد أن الإجابة الخاطئة التى تصدر عن الطفل فى الاختبارات المعتمدة على نموذج بياجيه تحتل أهمية خاصة كالإجابة الصحيحة فى فهم العمليات المعرفية لدى الطفل، وجوهر القياس هنا هو عملية حل المشكلة وليس نتائج هذا الحل.

وعلى الرغم من أن معظم الاختبارات من هذه الفئة لاتزال فى مرحلة البحث، إلا أنه توجد بعض الاختبارات تصلح للاستخدام العام نذكر منها:

١ – **مقاييس النمو العقلى:** من إعداد ليورنيدو Laurendeau وبينارد Pin-ard خلال الفترة ١٩٦٢ – ١٩٧٠ ضمن مشروع ضخم فى علم النفس النمو أجرى فى جامعة مونتريال بكندا ويقيس بعض مفاهيم بياجيه الأساسية وخاصة السببية (العلية) والمكان، وتطبق الاختبارات بطريقة الاستخبار الشفوى Oral questionnaires ـ كما يسميها الباحثان ـ والتى تقع فى منزلة بين طريقة بياجيه غير المقننة التى يسميها «الطريقة الكلينيكية» والطريقة المقننة الكاملة فى الاختبارات التقليدية فجميع الأسئلة مقننة، ولكن يمكن للفاحص أن يتخذ مسارات مختلفة بديلة من أجل مزيد من استطلاع واستكشاف عمليات التفكير لدى الطفل الفرد، وذلك اعتمادا على الاستجابات المبدئية التى تصدر عنه.

ويتم تصحيح البروتوكول الكامل لاستجابات الطفل لكل اختبار باعتباره وحدة وذلك فى ضوء المستوى الارتقائى الذى يدل عليه كيف الاستجابات، وقد أخضع الباحثان مقاييسهما للتحليل الإحصائى، فاختاروا عينة تقنين عددها ٢٥ طفلا، ٢٥ طفلة فى كل مستوى عمرى من سن ٢ – ٥ سنوات، وبلغ المجموع الكلى للعينة ٧٠٠ طفلا تمثل المجتمع الكندى المتحدث بالفرنسية فى مونتريال.

٢ – المقاييس الرتبية للنمو النفسى: أعد هذه المقاييس أوزجيرس Uzgiris, I.c. وهنت Huntعام ١٩٧٥ لأعمار أقل تصل إلى سن شهرين وتمتد إلى سن سنتين فقط، وهى مرحلة تشمل ما يسميه بياجيه المرحلة الحسية الحركية والتى تشتمل على ٦ مراحل (راجع وصفها بالتفصيل فى المرجع رقم (٢٤، ٣٤).

وتشمل البطارية ٦ اختبارات تقيس ما يلى: ثبات الشىء، تنمية وسائل للوصول إلى غايات بيئية مرغوبة، المبادأة، العلية الإجرائية، العلاقات بين الأشياء فى المكان، نمو المخططات. وقد حصل الباحثان على معامل اتفاق بين الملاحظين، ومعامل ثبات بإعادة الاختبار، لكل مقياس والتى تراوحت بين ٠,٨٠٢، ٠,٩١ كما حسب الباحثان الحد الأدنى لمعامل الرتبية Ordinality باستخدام طريقة جتمان المسماة Scologram فبلغ ٠,٥٠ إلا أن اناستازى (٦١: ٢٧٤) ترى أن تطبيق هذه الطريقة على المقاييس المعتمدة على نموذج بياجيه موضع خلاف لا يتسع المقام لتناوله بالتفصيل.

٣ – مجموعة تقريرالمفاهيم: وقد أعد هذه البطارية جولد شمد Goldschmid وبنتلر Bentler عام ١٩٦٨ للأعمار من ٤ – ٧ سنوات،وتقيس أحد المفاهيم المهمة فى نموذج بياجيه وهو مفهوم ثبات الخصائص Conservation (راجع ٢٤، ٣٤)، والذى يعد مؤشرا على تحول الطفل من مرحلة التفكير قبل العملياتى إلى التفكير باستخدام العمليات العيانية أو المحسوسة، وهو التحول الذى يحدد بياجيه حدوثه فى سن ٧ – ٨ سنوات.

ولهذا المقياس ثلاث صور، منها اثنتان متكافئتان فى قياس المكان الثنائى البعد والعدد والجوهر والكم المتصل والوزن والكم المنفصل، وقد تأكد امبيرقيا أنهما صورتان متكافئتان فى المتوسط والانحراف المعيارى، ومعامل الارتباط بينهما ٠,٩٥ أما الصورة الثالثة فنقيس المساحة والطول. ويتطلب تطبيق المقياس التزام الفاحص بتعليمات مكتوبة وتشمل عرض رسوم للمواد المستخدمة فى الاختبار وتوجيهات لمعالجتها وتناولها وتعليمات لفظية تقدم للمفحوصين.

وقد حسبت المعايير فى صورة مئينيات من أفراد عينة تتألف من ٥٠٠ طفل وطفلة تمتد أعمارهم بين ٤ – ٨ سنوات تمثل مدينة لوس أنجليس بولاية كاليفورنيا، وقد حسبت معاملات الثبات باستخدام ثبات المصححين ومعادلة كيورد ـ ريتشاردسون، وطريقة إعادة الاختبار، كما حسبت القياسية Scalability والرتبية للاختبار، وبالإضافة إل وصف بنيته باستخدام التحليل العاملى إلا أن العينات كانت صغيرة ويحتاج الاختبار لمزيد من البحث، ولعل أهم النتائج التى يزودنا بها المقياس هى قابليته للاستخدام فى البحوث النفسية المقارنة أو عبر الحضارية cross-cultural.

تعليق عام على المقاييس المعتمدة على نموذج بياجيه:

على الرغم من أن الاتجاه نحو بناء مقاييس نفسية تعتمد على نموذج بياجيه قد تأخر طويلا، إلا أنه يمثل فى السنوات الأخيرة أحد معالم التطور التى يشهدها علم النفس المعاصر، إلا أن ذلك يجب ألا يصرفنا عن بعض المشكلات المتضمنة فى هذه المقاييس ولعل أهمها، ظهور بعض مظاهر «عدم الاتساق» decalage فى التتابع المتوقع للنمو، ويتضمن هذا بعض الشك فى افتراضات الاتصال والانتظام والتتابع التى يطرحها نموذج بياجيه للنمو العقلى. فكثير ما يلاحظ أن المرحلة تطابق أداء الطفل تختلف من مهمة لأخرى، ليس فقط بالنسبة للعمليات المختلفة التى تطلب الحل، وإنما أيضاً حين تستخدم نفس العملية فى أنواع مختلفة من المحتوى.

إلا أن ما يجب أن ننبه إليه أن هذه المقاييس ارتبطت ارتباطات مرتفعة مع اختبارات الذكاء التقليدية المقننة، وترتبط بنفس الدرجة مع التحصيل المدرسى للأطفال فى الصف الأول، كما هو الحال بالنسبة لاختبارات الذكاء الفردية والجماعية، ومعنى ذلك أنه على الرغم من اختلاف المنهج فإن مهام بياجيه وأسئلة اختبارات الذكاء المقننة، وأسئلة الاختبارات التحصيلية بينها جميعا قدر من التداخل والاشتراك، وعلى الرغم من أن مهام بياجيه أصعب فى التطبيق وتستغرق وقتا أطول، إلا أن عائدها للتقويم النفسى كبير، فهى تقدم لنا وصفا كيفيا للنشاط العقلى للطفل لا توفره الأدوات الأخرى.

الإتجاهات المعاصرة فى تقدير* سلوك الأطفال فى مرحلة الطفولة المبكرة

اهتم علماء النفس فى العقد الأخير من القرن العشرين بقياس سلوك الأطفال فى مرحلة الطفولة المبكرة، وقد اعتمد تقدير Assessmentسلوك الأطفال فى هذه المرحلة على مقارنة أدائهم فى مهمة ما بالمعايير المرتبطة بذات السلوك للأطفال فى مثل أعمارهم.

وقد بذلت ومازالت تبذل – جهود كبيرة من المهتمين بتقدير سلوك الأطفال المعرفى وإمكانية التنبؤ به وذلك لأهمية اكتشاف مشكلات الضعف والتميز المعرفى حال وجودهما، وبالطبع يتطلب ذلك بناء الأدوات الملائمة لتحقيق الهدفين السابقين سواء فى حالة الضعف أو التميز.

ويعد أحد الاتجاهات المعاصرة قياس مهارات تجهيز المعلومات والذى اهتم به كل من (فاجن ١٩٩٢، فاجن/ ديترمان سنة ١٩٩٢) فى اناستازى وأوربانا سنة ١٩٩٧ : ٢٤٥) حيث ظهر هذا الاتجاه فى اختبارهم للذكاء Fagan Test of Infant Intelligence ، حيث اعتمد هذا الاختبار على اكتشاف الاطفال الذين يتميز سلوكهم بالجدة Novelityوالذى يمكن من خلال هذا السلوك دراسة قدرتهم على التجديد واسترجاع المعلومات، كما يمكن من خلال ذات المقياس التمييز بين الأطفال الأسوياء وأولئك الذين يعانون ولديهم مشكلات معرفية، وأيضا اكتشاف الأطفال الذين لديهم انتباه انتقائى للمثيرات التى تتميز بالجدة من سن ٣ شهور إلى ١٢ شهرا. والمثيرات المتضمنة فى الإختبار صور لوجوه انسان، وتقدر الدرجة بناء على مقدار الزمن الذى يظل الطفل ينظر فيها إلى الصورة التى تتميز بالجدة فى مقابل الصورة العادية (تقليدية) . وقد ثبت أن هذا الإختبار ينبىء الأداء فيه بالأداء فى اختبارات الذكاء، وقد بلغ معامل الإرتباط بين الأداء فى اختبار فاجن واختبارات الذكاء فى سن ٣ سنوات يتراوح بين ٠,٤٠، ٠,٦٠ .

وبالرغم من أن فكرة الإختبار جيدة إلا انه يتطلب الكثير من الجهد للحصول على البيانات التى تؤكد صدقه وقدرته على التمييز بين الأطفال الاسوياء وغير الاسوياء، لأن الدرجة التى يحصل عليها الطفل لا تحدد نقاط الضعف والقوة ولكنها يمكن أن تفيد فى تصنيف الأطفال الى فئات، كما أن تقدير أداء الاطفال والكبار على حد سواء يتطلب طرقا متعددة ومصادر متعددة للمعلومات عن المجالات وثيقة الصلة باللغة والحركة والمهارات الاجتماعية. ولمواجهة هذا المطلب فقد بذلت جهود وتطورت بعض النظم فى التقويم من المتخصصين فى المجالات المختلفة.

وتشير اناستازى وأوربانا (١٩٩٧ :٢٤٥) أن تقدير السلوك المتعدد التخصصات لسلوك الأطفال فى مرحلة الطفولة المبكرة يعد اطارا مرجعى فى التعرف على «الأطفال فى خطر» Children at risk فى التأخر فى النمو من الميلاد وحتى ٣ سنوات، وأطلق على القياس متعدد التخصص The Infant - Todler Develop-mental Assessment (IDA) وتساعد المادة التى يتم الحصول عليها من هذا القياس تجميع وتفسير وتركيب المعلومات التى تم جمعها من خلال خطة دراسة نمو الطفل، وتتضمن هذه الخطة اشراك الآباء فى كل مرحلة، واستعراض تقرير طبى عن الطفل،

كما يتضمن تقريرا عن نمو الطفل من خلال الملاحظة للطفل والمقابلة الشخصية للآباء ورعاة الأطفال (Care givers).

ويتكافىء هذا المقياس مع مقاييس نمو الأطفال من الميلاد وبروفيل النمو حتى ثلاث سنوات، وبالرغم من أن بروفيل النمو يتطلب تطبيقا معياريا وطريقة مقننة للتصحيح لثمانية مجالات للنمو وتطوره، فإن هذا التقدير ينسب إلى العمر الزمنى وما يرتبط به من خصائص أكثر من اعتماده على معايير أو مقارنات . وعلى ذلك فإن مستخدم الاختبار يجب ان يكون على دراية تامة بالمقياس ، كما يحتاج أيضا دراسات وبحوث لتقنين فعاليته فى المجالات الكلينيكية.

وبالرغم من أهمية الاتجاهات المعاصرة والتى تتطلب تقديرا متكاملا وشاملا لسلوك الأطفال فى مرحة الطفولة المبكرة، فإنه يجب التأكيد على التأثيرات البيئية والإقتصادية على أداء الطفل، مع أهمية دراسة تأثير العلاقة بين تقدير سلوك الطفل والتدخل العلاجى حيث ان التشخيص يرتبط ارتباطا تاما مع العلاج. وعلى ذلك يجب ان تكون هناك ارشادات موجهة للمعلمين والآباء والرعاة تعتمد على البروفيل الشخصى للطفل متضمنة جوانب التميز والصعوبات والتى يوصى بأهمية التدخل المبكر فيهما سواء بالإثراء أو العلاج.

جهود المركز القومى للإمتحانات والتقويم التربوى فى تحديد معايير نمو طفل ما قبل المدرسة

قام فريق من المركز القومى للإمتحانات والتقويم التربوى برئاسة أ.د. محمد عماد الدين اسماعيل بدراسة هامة هدفت الى التوصل الى معايير مصرية لنمو طفل ما قبل المدرسة (٢-٦ سنوات)، كما هدفت الدراسة إلى اعداد اداة قياس مبسطة يسهل تطبيقها (الوالدان - الطبيب - المعلم- الباحث) لتمكينهم من تقدير مستوى نمو الطفل، كما تساعد على الإكتشاف المبكر للأطفال متأخرى النمو والأطفال المعرضين للإعاقة وكذلك الأطفال المتفوقين.

وقد أجريت الدراسة على عينة (٨١٤) طفلا وطفلة من القاهرة الكبرى ممثلة للشرائح الإقتصادية والاجتماعية والثقافية المختلفة، وتراوحت أعمار العينة من ٢-٦ سنوات. وقد اعتمدت الدراسة على نموذج «راش»، وبرنامج مايكرو سكيل - Micro

scale فى بناء مقاييس النمو . وقد توصلت الدراسة الى «وحدة» لتدريج كل من صعوبة المفردات وقدرات الأطفال وأطلق عليها «وحدة منف» حيث تؤدى الى تدرج مئوى يسهل معه التفسير. وقد تم التوصل الى مقاييس للنمو النفسى للطفل مدرجة بطريقة «راش» وهى:

١- مقياس الحركات الكبيرة.

٢- مقياس الحركات الدقيقة.

٣- مقياس ما قبل الكتابة.

٤- مقياس مساعدة الذات.

٥- مقياس المجال المعرفى.

٦- مقياس اللغة.

٧- مقياس السلوك الاجتماعى.

وأمكن من خلال هذه المقاييس التوصل إلى معايير دقيقة وموضوعية لنمو الطفل المعيارى ومقارنته بأقرانه وتحديد مستواه بينهم.

الاختبارات العملية للذكاء.

توجد طائفة من المقاييس العقلية تقيس الأداء العملى ولا تتضمن أى اختبارات لفظية أو لغوية، وقد أشرنا إلى أن مقاييس وكسلر تضمنت قسما عمليا أدائيا بالإضافة إلى القسم اللفظى اللغوى، وبالفعل فإن المقاييس العملية للذكاء استفادت من التطورات التى حدثت فى ميدان القياس النفسى للذكاء استفادت من التطورات التى حدثت فى ميدان القياس النفسى بعد ظهور مقياس وكسلر- بليفو فى ميدان قياس الذكاء، بالإضافة إلى الجهود التى بذلت لقياس الذكاء، «العملى»، ولهذا يمكن القول أنه توجد فئة من اختبارات الذكاء العام تعتمد اعتمادا كليا على الأداء العملى وحده وتتطلب نتيجة لذلك تطبيقا فرديا.

وقد يكون أول مقياس عملى للذكاء هو لوحة الأشكال التى ابتكرها سيجان Se-guin عام ١٨٦٦ لتستخدم فى قياس ذكاء ضعاف العقول، وشاعت هذه اللوحة بعد

ذلك ، بل وأدخلت فى كثير من اختبارات الذكاء العملية اللاحقة، وتتألف فى الوقت الحاضر من ١٠ قطع خشبية توضع أمام الطفل بترتيب مقدن، ثم يطلب منه وضع كل قطعة فى موضعها من اللوحة بأسرع ما يمكن، ويسمح له بثلاث محاولات لكل قطعة، ودرجة المفحوص لكل قطعة هى أسرع زمن يضع فيه القطعة الخشبية من بين المحاولات الثلاثة.

وفى عام ١٩١٤ أعد نوكس سلسلة من الاختبارات العملية تطبق دون استخدام اللغة، تضمنت لوحات أشكال متزايدة الصعوبة، بالإضافة إلى اختبارين أحدهما يسمى اختبار السفينة، والثانى يسمى اختبار المكعبات، ويتألف الأول من ١٠ قطع خشبية مستطيلة الشكل ترتب فى إطار خشبى لتكون فى النهاية صورة لسفينة فى البحر، أما الثانى فهو اختبار للذاكرة المباشرة لسلسلة من الحركات تصدر عن الاخص فى صورة نقر محدد تحديدا مسبقا بطريقة مقننة على كل مكعب من المكعبات الأربعة التى يتألف منها الاختبار ثم يطلب من المفحوص أن يحاكيه.

وقد يكون مقياس بنتر - بيترسون Pinter-Patrson الذى ظهر عام ١٩١٧ أكثر الاختبارات أهمية من حيث التقنين، وهو يتألف من مجموعة من الاختبارات مشتقة من الاختبارات العملية السابقة، وخاصة اختبارات سيجان وهيلى ونوكس وغيرها، ويبلغ عددها الكلى ١٥ اختبارا (وتوجد صورة مختصرة للمقياس تتألف من ١٠ اختبارات فقط).

وفى عام ١٩٢٣ ظهر اختبار كوهس لرسم المكعبات Kohs Block Design Test كمقياس يصاحب مقياس ستانفورد - بينيه فى التعرف على المتخلفين عقليا. أو يحل محله، ويتألف من مجموعة من المكعبات المتماثلة الملونة سطوحها الست بالألوان الآتية: الأحمر، الأزرق، الأصفر، الأبيض، الأصفر والأبيض معا، الأحمر والأبيض معا، وتعرض على المفحوص بطاقة - فى كل مرة - عليها رسم ملون ويطلب منه إنتاجه عن طريق تجميع المكعبات المناسبة، ويختلف عدد المكعبات المطلوب لمختلف الرسوم تبعا لمستويات الصعوبة من ٥ مكعبات إلى ١٦ مكعبا، وتعتمد الدرجة على زمن الاستجابة وصوابها.

ومن الاختبارات العملية للذكاء التى ظهرت فى الربع الأول من القرن الماضى

وذاعت شهرتها اختبارات متاهات بورتيوس Porteus التى ظهرت لأول مرة عام ١٩١٤، وتتألف من سلسلة من المتاهات المرسومة على الورق مرتبة حسب الصعوبة ويمكن إعطاؤها دون تعليمات لفظية وذلك باستخدام المتاهات السهلة على سبيل التمثيل، وتصلح هذه الاختبارات للمستويات العمرية من ٣ سنوات حتى مستوى الرشد، وفيها يطلب من المفحوص أن يتتبع بالقلم الرصاص أقصر طريق من نقطة البداية إلى نقطة النهاية فى متاهة الورق، دون رفع القلم مطلقا من على الورقة، وحين يقع المفحوص فى «خطأ» سواء خرج عن الطريق أو دخل فى ممر مغلق يوقفه الفاحص عن لأداء ويعطيه محاولة أخرى على متاهة مماثلة، فإذا أخطأ فى المحاولة الثانية يعطى درجة «فشل» فى هذا المستوى، ويسمح للمفحوص فى المستويات العليا بأربع محاولات، وتعتمد الدرجة على الزمن المستغرق، ويوضع فى الاعتبار أثناء التصحيح رقم المحاولة التى اجتاز فيها المفحوص المتاهة بنجاح، ويؤكد بورتيوس أن اختباراته تقيس قدرات الاستبصار والتخطيط التى تعجز الاختبارات اللفظية عن قياسها.

ويمكن القول أن أكثر بطاريات اختبارات الذكاء العملية شيوعا فى الوقت الحاضر بطاريتان: أولاهما ظهرت فى الولايات المتحدة عام ١٩٣٠ وتسمى اختبارات آرثر العملية وثانيهما اختبارات الكسندر.

وتتألف بطاريات آرثر Arthur Point Scale of Performance Tests من أفضل الاختبارات العملية التى ظهرت قبل ذلك، وتتكون من ٩ اختبارات هى:

١ – مكعبات نوكس.
٢ – لوحة الأشكال لسيجان.
٣ – متاهات بورتيوس.
٤ – اختبار كوهس لرسم المكعبات.
٥ – لوحة الأشكال التى تتألف من مربع وصليب ينقسم كل منها إلى ٤ أقسام توضع فيها أجزاؤها.
٦ – لوحة الأشكال المتشابهة: وهى أكثر صعوبة لأن القطع الخشبية فيها أقرب إلى التشابه.
٧ – المانيكان ويتألف من شكل خشبى لرجل تجمع أجزاؤه: الذراعان والساقان

والرأس والجذع. والبروفيل ويتألف من قطع خشبية تجمع معا لتكون وجه إنسان فى صورة بروفيل.

٨ – تكملة الصور (سهل).

٩ – تكملة الصور لهيلى Healy وهو أصعب من الاختبار السابق لأن القطع الخشبية فيه متماثلة.

وفى عام ١٩٤٧ ظهرت طبعة معدلة من مقياس آرثر الأدائى تتألف من ٥ اختبارات فقط تشمل تعديلات على اختبارات مكعبات نوكس ولوحة الأشكال لسيجان ومتاهات بورتيوس، بالإضافة إلى اختبار آخر لتكملة الصور مأخوذ من سلسلة هيلى - فرنالد، واختبار جديد يشبه اختبار كوهس لرسم المكعبات فى أنه يطلب من المفحوص محاكاة رسوم متزايدة الصعوبة معروضة أمامه على بطاقات، ويختلف عنه فى أن الرسوم مصنوعة من الورق المقوى وملونة بألوان عديدة. ويسمى هذا الاختبار الجديد باختبار آرثر لرسوم الاستلسل.

وفى عام ١٩٣٦ ظهر فى بريطانيا اختبار مماثل هو مقياس كولنز - دريفر Col-lins-Drever ويتألف من ١١ اختبارا عمليا، يضم الاختبارات التسع آنفة الذكر فى مقياس آرثر (مع اعتبار البروفيل اختبارا منفصلا، فتصبح ١٠ اختبارات). بالإضافة إلى اختبار لصورة نجمة خشبية مقسمة إلى أجزاء يقوم المفحوص بتركيبها، و(توجد صورة أخرى له تتأف من ٧ اختبارات وتصلح للأطفال الصغار).

أما الاختبارالهام الثانى هو الاختبار الأدائى لألكسندر Alexander وهو محصلة الدراسات العاملية التى قام بها العالم الإنجليزى الكسندر ويتألف من ٣ اختبارات، اثنان منها هما اختبارا مكعبات كوهس وبناء المكعبات اللذان أشرنا إليهما، بعد أن عدل الكسندر إجراءات تصحيحهما بحيث تسمح بإعطاء أكثر من درجة واحدة للنجاح فى رسم النماذج المختلفة وللزمن المستغرق، أما الاختبار الثالث فهو اختبار جديد أطلق عليه الكسندر اسم اختبارالإزاحة Passalong، ويتألف من ٩ صناديق صغيرة كل منها يحتوى على قطعة خشبية واحدة لونها أحمر وعدة قطع لونها أزرق (هذه القطع قد تكون على شكل مربعات أو مستطيلات)، وعن طريق تحريك القطع بالانزلاق، دائريا تنتقل القطعة الحمراء من أسفل إلى أعلى) دون رفع القطع من أماكنها مطلقا).

وقد رتبت مشكلات الاختبارتبعا لمستويات الصعوبة، وتصحح حسب الزمن المستغرق حتى الوصول إلى الحل الصحيح.

وتحول درجات هذاالمقياس إلى أعمار عقلية، وقد أعدت له معايير للذكور والإناث (الإنجليز) من سن ٧ سنوات إلى ١٩ سنة، وتسمى نسبة الذكاء الناتجة «نسبة القدرة العملية Practical Ability Ratio وتصلح فى رأى الكسندر فى انتقاء الطلاب للتعليم الفنى.

وبالطبع فإن معظم الاختبارات التى تستخدم فى قياس ذكاء المعوقين تتخذ فى جوهرها صورة الاختبارات العملية، ومن ذلك اختبار هيسكى ـ نبراسكا Hiskey-Nebraska لقياس ذكاء الصم الذى يتألف من ١٢ اختبارا فرعيا هى:

١ – أنماط الخرز
٢ – ذاكرة الألوان
٣ – تعيين الصور
٤ – تداعى الصور
٥ – ثنى الورق (أنماط)
٦ – مدى الانتباه البصرى .
٧ – أنماط المكعبات
٨ – إكمال الرسوم
٩ – ألغاز المكعبات
١٠ – المتشابهات المصورة
١١ – الاستدلال المكانى
١٢ – ذاكرة الأرقام.

أما بالنسبة للمكفوفين قد ظهرت صور معدلة من مقياس ستانفورد ـ بينيه ومقاييس وكسلر تلائم ظروف إعاقتهم.

الفصل الثامن
الاختبارات الجماعية للذكاء

أهمية الاختبارات الجماعية للذكاء:

ينتمى مقياسا ستانفورد - بينيه ووكسلر إلى فئة الاختبارات الفردية للذكاء، وهى أصلح ما تكون للأغراض الكلينيكية، كما أشرنا فى الفصل السابق. أما فى برامج التقويم النفسى التى تتطلب اختبار أعداد كبيرة من المفحوصين فى الميادين التربوية والصناعية والإدارية والعسكرية تصبح هذه الاختبارات الفردية أكثر تكلفة فى الجهد والمال والوقت ولذلك ظهرت الاختبارات الجماعية لمواجهة مثل هذه الصعوبات وبخاصة فى الحرب العالمية الأولى حين أعد الأخصائيون النفسيون العسكريون فى الولايات المتحدة اختبارا ألفا (وهو اختبار لفظى) واختبار بيتا (وهو اختبار غير لفظى) لأغراض الاستخدام الحربى.

خصائص الاختبارات الجماعية:

لا تتطلب الاختبارات الجماعية فى معظم الحالات فاحصا مدربا تدريبا خاصا، كما أن الإجابات فيها تسجل على أوراق منفصلة للإجابة، ولهذا لا تتطلب هذا النوع من العلاقات الإنسانية المباشرة بين فاحص ومفحوص.

وللاختبارات الجماعية ميزات مهمة منها تيسير إجراءات التطبيق، واتساع حجم العينات التى يمكن أن تطبق عليها هذه الاختبارات وحسن تمثيلها للأصل الإحصائى السكانى العام، ويسر وسائل التصحيح الموضوعية وتوافر معايير أكثر دقة، وشمول عينات التقنين التى تصل إلى أكثر من ربع مليون فى بعض الاختبارات الجماعية فى حين أنها لم تتعد ٤٠٠٠ مفحوص فى أحسن حالات الاختبارات الفردية.

إلا أن لهذه الاختبارات مشكلاتها أيضاً، ومن ذلك ضعف العلاقة بين الفاحص والمفحوص، وصعوبة تأكد الفاحص من تعاون المفحوص، وعدم إحاطته بدوافع المفحوص لأخذ الاختبار أوميله واهتمامه به. بل يصعب على الفاحص أن يتعرف على المشكلات الطارئة على موقف الاختبار وخاصة بالنسبة للأحوال الداخلية للمفحوص أثناء الاختبار الجماعى، وكذلك فإن بعض المفحوصين يواجهون صعوبات

فى موقف الاختبار الجماعى، وبالتالى يتعطل أداؤهم عنه فى موقف الاختبار الفردى، كما توجد أدلة على أن الأشخاص المضطربين انفعاليا يؤدون فى مواقف الاختبار الفردى أفضل من الاختبار الجماعى، وأخيراً إن الاختبارات الفردية، لا تيسر للفاحص الحصول على ملاحظات إضافية عن سلوك المفحوص أوتحديد العوامل المسئولة عن ضعف الأداء فى بعض الأسئلة أو بعض الاختبارات الفرعية.

وتختلف الاختبارات الجماعية عن الاختبارات الفردية فى القالب وفى طريقة ترتيب الأسئلة، فأغلب الاختبارات الجماعية الحديثة تعتمد على أسئلة الاختيار من متعدد، وقليل منها يستخدم أسئلة النهايات المفتوحة، كما تصنف الأسئلة المتشابهة فيها إلى اختبارات فرعية منفصلة توحيدا للتعليمات. وفى كل اختبار فرعى ترتب الأسئلة عادة تبعا لمستويات الصعوبة.

وقد تنشأ بعض الصعوبات العملية من استخدام نظام الاختبارات الفرعية المنفصلة، ومن ذلك أن الفاحص الأقل خبرة والأقل دقة قد يقع فى أخطاء ضبط زمن الاختبار وخاصة فى الاختبارات ذات الزمن القصير. وحتى يمكن تحقيق قدر من التوازن بين استخدام الزمن المعد للاختبار وتنظيم الأسئلة عل نحو يسمح لجميع المفحوصين بمحاولة الإجابة على جميع أنواعها بمستويات صعوبة متزايدة يلجأ بناة الاختبارات النفسية فى الوقت الحاضر إلى النظام الحلزونى الشامل Spiral-Omnibus وفيه تعرض أولا أسهل الأسئلة من كل نوع ثم يتلوها الأصعب من كل نوع فالأصعب، وهكذا على نسق حلزونى تصاعدى فى مستوى الصعوبة، وحتى يمكن تجنب الحاجة إلى إعادة التعليمات بالنسبة لكل سؤال ولخفض عدد مرات التحول فى التأهب الناتج عن التعليمات (٢٤) يلجأ بعض مؤلفى الاختبارات إلى استخدام النظام الحلزونى الشامل بالنسبة لمجموعات من الأسئلة، يتألف كل منها من ٥ – ١٠ أسئلة.

ويرى بعض نقاد الاختبارات الجماعية أن هذا النوع يفرض قيودا على استجابات المفحوص ، وقد وجه هذا النقد خاصة إلى أسئلة الاختيار من متعدد وإلى بعض الأنواع الشائعة من الاسئلة مثل القياس التخيلى والمتشابهات والتصنيف بل إن بعض هذه الأسئلة - فى رأى هؤلاءالنقاد - تعاقب المفحوص المبتكر الذى يدرك تضمينات غير عادية فى السؤال، ويرد أنصار هذه الاختبارات على هذا النقد بأن هذه الحالات

نادرة الحدوث وخاصة إذا كانت أسئلة الاختبار حللت تحليلا دقيقا وتوافرت عنها بيانات صدق كافية.

وفى رأى فريق آخر من النقاد أن الاختبارات الجماعية لاتوفر للفاحص فرصاً لتحليل الأخطاء والبحث عن أسباب اختيار المفحوص لإجابة مَعينة، وهذا النقد صحيح بالطبع، ولهذا السبب وغيره لابد من أن يطبق مع اختباره الجماعى اختبارا فرديا على المفحوص أو أن يحصل على معلومات إضافية عنه من مصادر أخرى وخاصة فى الحالات التى تتطلب اتخاذ قرارات مهمة بالنسبة له. وخاصة فى الحالات الهامشية التى يصعب اتخاذ قرار واضح بشأنها.

وتوجد مشكلة أخرى بالنسبة لهذا النوع من الاختبارات وهى عدم المرونة التى تتمثل فى أن كل مفحوص يختبر فى جميع الأسئلة، وفى هذا استهلاك للوقت دون مبرر، والأجدى أن يركز كل مفحوص على الأسئلة المناسبة لمستوى قدرته، وإذا حدث ذلك فإن المفحوص يتجنب الملل الناجم عن العمل مع أسئلة شديدة السهولة، أو الشعور بالإحباط والقلق الناجم عن محاولة الإجابة على أسئلة شديدة الصعوبة وتتجاوز مستوى قدرته، والاختبارات الفردية فى هذا الصدد لها ميزة واضحة، ولعلنا نذكر أنه فى مقياس ستانفورد - بينيه يتم انتقاء الأسئلة التى تقدم للمفحوص على أساس استجاباته السابقة، وتعطى له الأسئلة التى تقع فى نطاق مستوى قدرته.

دور الحاسب الالكترونى فى بناء نظام اختبارى تكيفى:

بذلت فى السنوات الأخيرة محاولات للجمع بين مزايا كل من الاختبارات الفردية والجماعية، وخاصة فى مجال المواءمة بين اسئلة الاختبار وخصائص استجابات المفحوص الفرد، ويسمى هذا الأسلوب بأسماء مختلفة منها النظام الاختبارى التكيفى adaptive أوالتتابعى sequential أو المتفرع branched أو التفصيل tailored أوالتفريدى individulized أو المبرمج programmed أو الدينامى dynamic أوالمشروط بالاستجابة response-contingent (راجع أنستازى ٦١ : ٣٠١ - ٤٠٣)، وهذه الأساليب أيسر فى التناول حين تستخدم الحاسب الألكترونى فى تطبيق الاختبار.

وتتلخص طريقة بناء النظام التكيفى (أو غير ذلك من الصفات السابقة) للاختبار فى البدء بإعطاء اختبار قصير يسمى الاختبار الاستكشافى routing rest تشمل أسئلته

مدى واسعا من الصعوبة، وفى ضوء استجابة المفحوص فى هذا الاختبار يوجه إلى اختبار من بين عدة اختبارات ذات مستويات صعوبة مختلفة، عادة ما يكون أطول الاختبار الاستكشافى إلا أنه يتفق مع مستوى قدرة المفحوص وتسمى هذه الطريقة «طريقة المرحلتين».

وتوجد طريقة أخرى لبناء هذا النظام الاختبارى تسمى «طريقة الهرم» وفيها يبدأ جميع المفحوصين بالإجابة على سؤال متوسط الصعوبة، فإذا كانت إجابة المفحوص عليه صحيحة يوجه إلى سؤال أكثر صعوبة، أما إذا كانت إجابته خاطئة يوجه إلى سؤال أسهل ويتكرر ذلك بعد كل استجابة على كل سؤال.

وقد استخدمت هاتان الطريقتان فى بناء اختبارات من نوع الورقة والقلم المعتادة أو من النوع الذى يتطلب استخدام الحاسب الالكترونى، أما الطرق الأكثر تعقيدا والتى لا تتطلب انماطا ثابتة محددة مقدما من تتابع الأسئلة فلايمكن إجراؤها إلا بالحاسب الالكترونى.

نظرية السمة الكامنة:

تستخدم الإجراءات المعتمدة على الكمبيوتر فى تطبيق الاختبار نظرية الاستجابة على المفردة item response theory وهى التى تسمى نظرية السمة الكامنة latent trait theory فى بعض الأحيان، وتقوم هذه النظرية على خاصتين أساسيتين لمنحنى المفردة أو السؤال وهما:

١ - **مقياس القدرة:** ويقدر بالعدد الكلى للاستجابات على الاختبار وغير ذلك من المعلومات التى تتوافر للباحث عن استجابات عينة معينة من المفحوصين على الاختبار.

٢ - **احتمال حدوث الاستجابة الصحيحة لمفردة معينة:** وهو دالة لوضع المفحوص فى مقياس القدرة. ويحسب هذا الاحتمال من البيانات التى تمثل نسب الذين يجتازون السؤال أو المفردة من الأفراد ذوى المستويات المختلفة فى القدرة.

ويتألف النموذج الكامل لنظرية السمة الكامنة من ثلاثة بارامترات تشتق رياضيا من البيانات الامبريقية عن مفردات الاختبار وهى:

١ - **تمييز المفردة:** ويتحدد رياضيا بميل slope المنحنى الدال على علاقة مقياس القدرة باحتمال حدوث الاستجابة الصحيحة للمفردة، وهما اللذان يؤلفان المحور الأفقى والمحور الرأسى على التوالى فى هذا المنحنى، ويرتبط هذا البارامتر عكسيا بالمسافة المطلوبة فى متصل مقياس القدرة حتى يزداد احتمال حدوث الاستجابة الصحيحة للسؤال، وكلما ازدادت قيمة التمييز يزداد الميل انحدارا.

٢ - **صعوبة المفردة:** ويتطابق هذا البارامتر مع موضع المفردة فى محور مقياس القدرة حين يكون احتمال الاستجابة الصحيحة ٠,٥٠.

٣ - **التخمين:** وهذا البارامتر يستخدم فى أسئلة الاختيار من متعدد والتى من أجلها صممت خصيصاً نظرية السمة الكامنة، ويمثل احتمال حدوث الاستجابة الصحيحة بمحض الصدفة ويتحدد بالخط المقارب الأدنى lower asymptote فى المنحنى الخاص بالمفردة (٦١:١٩١).

ويعتمد الحاسب الالكترونى على الخصائص السابقة التى تتوافر فى مجموعة كبيرة من المفردات، فبعد أن يستجيب المفحوص على سؤال له خصائص معينة ينتقى الكومبيوتر السؤال التالى على أساس الاستجابات السابقة التى صدرت عن المفحوص لخفض أثر أخطاء التقدير، ويمكن أيضاً أن تستخدم هذه الطريقة للمواءمة بين نقطة البداية فى الاختبار أو السؤال المبدئى فيه ومستوى قدرة المفحوص، وفى هذه الحالة يعتمد الحاسب الالكترونى على البيانات السابقة المتوافرة عن المفحوص مثل نتائجه فى الاختبارات الأخرى أو فى الامتحانات المدرسية أو المعلومات التى يقدمها المفحوص عن نفسه فى بداية جلسة الاختبار.

وتعتمد درجة المفحوص فى هذه الحالة ليس على الأسئلة التى يجيب عنها إجابة صحيحة، وإنما على مستوى الصعوبة والخصائص السيكومترية الأخرى للأسئلة.

بطاريات اختبارات الذكاء المتعددة المستويات:

تظهر مقارناتنا السابقة بين الاختبارات الجماعية من ناحية الاختبارات الفردية والاختبارات المعتمدة على الحاسب الالكترونى من ناحية أخرى أن الاختبارات الجماعية للذكاء تقدم لجميع المفحوصين نفس المجموعة من الأسئلة بصرف النظر عن استجاباتهم الفردية، ولهذا السبب نلاحظ أن أى اختبار جماعى يجب أن يشمل

مدى محدود نسبيا من الصعوبة بحيث يلائم مستوى معين من العمر أو الصف الدراسى أو القدرة. وحتى يمكن توافر مقاييس للنمو العقلى لمدى واسع نسبيا قام الباحثون فى هذا الميدان ببناء بطاريات متعددة المستويات multilevel . وبهذه الطريقة يتم اختبار المفحوصين فى المستوى الملائم لهم، ويمكن للمستويات الأخرى أن تستخدم لإعادة اختبار نفس المفحوصين فى مراحل تالية، أو فى أغراض التقويم النفسى المقارن بين مختلف الأعمار أو الصفوف الدراسية، ويتوافر فى هذه البطاريات أرضية وسقف مناسبان للأفراد الذين يقعون على أطراف التوزيع سواء بالنسبة للعمر أو الصف الدراسى وذلك نتيجة التداخل بين البطاريات المتتابعة، إلا أننا يجب أن ننبه هنا إلى أن المواءمة بين صعوبة المفردات ومستوى قدرة المفحوصين ليست بمستوى الدقة الذى يتوافر فى الاختبارات الفردية أو الاختبارات المعتمدة على الكمبيوتر، وإنما عادة ما تكون هذه المواءمة تقريبية وعادة ما تتم فى ضوء المعلومات العقلية التى تتوافر لدى الفاحص عن المفحوص وخاصة عمره الزمنى أو صفه الدراسى وليس فى ضوء استجاباته على أسئلة الاختبار.

وأفضل مجال لتطبيق الاختبارات المتعددة المستويات هو المدرسة حيث تزداد الحاجة إلى المقارنة بين المفحوصين خلال سنوات عديدة، ولهذا السبب فإن معظم المستويات فى هذه البطاريات تتحدد فى ضوء الصفوف الدراسية، ويتوافر فى مستويات هذه البطاريات درجة مناسبة من المتصل سواء من حيث المحتوى أو العمليات النفسية والعقلية المقيسة، كما أن الدرجات التى يحصل عليها المفحوص من مختلف المستويات تتخذ نفس الصورة. بالإضافة إلى أن عينات التقنين فى هذه البطاريات تكون فى العادة متكافئة على نحو لايتوافر فى الاختبارات التى تقنن منفصلة. وأخيراً فإن بعض هذه البطاريات قد يمتد على نطاق واسع ابتداء من مرحلة الحضانة ورياض الأطفال حتى مستوى القبول فى الجامعة.

وتصاغ معايير معظم هذه البطاريات فى صورة نسب ذكاء انحرافية أو ما يشبهها من درجات معيارية معدلة، وبعضها يستخدم عدة أنواع من المعايير كالمئينيات والمعايير السباعية ومعايير الفرق الدراسية، وتعطى اختبارات هذه البطاريات أنواعا مختلفة من الدرجات ومن ذلك الدرجة الكلية فى البطارية بالإضافة إلى درجات منفصلة للاختبارات اللفظية والكمية، أو الأقسام اللغوية وغير اللغوية فى البطارية.

التقويم النفسى ـــــــ

تتطلب من الطفل معرفة باللغة المنطوقة، ولذلك فإنها تصف محتوى الاختبار ولا تصف سلوك المفحوص، فاختبارات الفهم اللفظى يمكن أن تطبق فى هذا السن باستخدام الصور، وتقاس مفردات الطفل أو فهمه للجمل بهذه الطريقة، وتفضل أنستازى (٦١ : ٢١٩) ـ منعا للخلط ـ أن تسمى هذه الاختبارات : الاختبارات «المصورة» pictorial أو الاختبارات التى لا تعتمد على القراءة non-reading بدلا من عبارة الاختبارات غير اللفظية.

ومن أمثلة الاختبارات التى تصلح لهذا المستوى اختبار القدرة العقلية لأوتس ولينون Otis-Lennon (المستوى الأولى) الذى ظهر طبعته الأخيرة عام ١٩٧٩ – ١٩٨٠، وله صورتان (ج. ك) لكل المستويات التى يصلح لها هذا الاختبار (وتمتد إلى نهاية المرحلة الثانوية) ويتكون المستوى الأولى Primary level من مستويين فى الواقع : المستوى ١ (لأطفال الحضانة) والمستوى ٢ (للنصف الأول من الصف الابتدائى الأول)، وهذان المستويان متطابقان فى المحتوى، ولايختلفان إلا فى طريقة تعبير الطفل عن استجابته، ففى مستوى الحضانة يضع الطفل دائرة حول الإجابة الصحيحة، أما فى المستوى الثانى فإن الطفل يسود بالقلم الرصاص مكان الإجابة الصحيحة فى ورقة الإجابة.

وتعطى تعليمات الاختبار شفويا، كما أن لكل سؤال زمنا محددا (حوالى ١٥ ثانية) ويتطلب الاختبار ككل من ٢٥ – ٣٠ دقيقة، ويطبق على جلستين بينهما فترة راحة قصيرة، فى كل جلسة يعطى قسم من قسمى الاختبار، ويتكون القسم الأول من ٢٣ سؤالا من أسئلة التصنيف وفيها يطلب من الطفل أن يضع علامة تحت الصورة (أو الشكل الهندسى) الذى لا ينتمى إلى فئة من ٤ صور (أو رسوم هندسية)، أما القسم الثانى فيتكون من ٣٢ سؤالا تقيس التصور الذهنى اللفظى Verbal Conceptualixation، والاستدلال الكمى quantitative reasoning والمعلومات العامة، والقدرة على اتباع التعليمات.

ومن أمثلة التصور الذهنى اللفظى أن يطلب من الطفل وضع علامة تحت الصورة التى تدل على نار، وفى الاستدلال الكمى يطلب منه مثلا وضع علامة تحت عدد النقط الذى يساوى عدد الأقسام المعطاة لدائرة مرسومة. وفى المعلومات العامة

يطلب منه وضع علامة تحت صورة الشيء الذى نستخدمه فى الحديث مع الآخرين (تليفون) . وفى اختبار اتباع المعلومات يطلب من الطفل مثلا وضع علامة تحت الصورة التى تدل على زجاجة داخل مربع وعليها علامة نجمة، ويحصل المفحوص فى هذا الاختبار بمستوياته المختلفة على مؤشر للقدرة المدرسية.

وقد قنن هذا المقياس فى أمريكا (كما قننت مقاييس المستويات الأعلى التى يقيسها نفس الاختبار) على عينة كبيرة ممثلة للمجتمع الأصلى فى أمريكا، ووضعت معاييره فى صورة نسب ذكاء انحرافية، متوسطها ١٠٠ وانحرافها المعيارى ١٦ . كما أن له معايير مئينية وتساعية تتحدد فى ضوء فئة العمر والصف الدراسى، وللاختبار معامل ثبات مرتفع باستخدام طريقة كيودر- ريتشاردسون وطريقة إعادة الاختبار، كما أن له صدق تلازمى وصدق تنبؤى جيد باستخدام درجات الاختبارات التحصيلية كمحك.

ويمكن استخدام هذه الاختبارات فى قياس قدرات الاستدلال المجرد عند الأطفال من ذوى صعوبات القراءة أوالتعويق التربوى من الصفوف الدراسية الأعلى.

اختبارات ذكاء تلاميذ مرحلة التعليم الأساسى:

من أشهر الاختبارات الجماعية التى تصلح ابتداء من الصف الرابع بمرحلة التعليم الأساسى المستوى (ب) من بطارية لورج- ثورنديك المتعددة المستويات Lorge-Thorndike Multi-Level Battery والتى ظهرت عام ١٩٦٤ . ويصلح هذا المستوى ابتداء من الصف الرابع الابتدائى، ولهذا الاختبار صورتان يتضمن كل منهما ٥ اختبارات لفظية و٤ اختبارات غير لفظية، وتشمل الاختبارات اللفظية المفردات، وتكملة الجمل والإستدلال الحسابى والتصنيف اللفظى والمتماثلات، أما الاختبارات غير اللفظية فتشمل: تصنيف الأشكال، وسلاسل الأعداد ومتماثلات الأشكال، ويتطلب الاختبار حوالى ٣٥ دقيقة للقسم اللفظى. و٣٧ دقيقة للقسم غير اللفظى، ورغم أن لكل اختبار فرعى زمنه المحدد، إلا أن هذه الاختبارات توصف بأنها من اختبارات القوة.

ومن البطاريات المتعددة المستويات التى لها أهمية خاصة اختبار القدرات المعرفية Cognitive Abilities Test من إعداد لورج وثورنديك وتتألف هذه البطارية الكلية من مستويين أوليين (ابتداء من مرحلة الحضانة ورياض الأطفال

وحتى الصف الثالث من مرحلة التعليم الأساسى فى مصر) ومن طبعة متعددة المستويات تشمل ابتداء من الصف الثالث من مرحلة التعليم الأساسى وحتى نهاية المرحلة الثانوية، ويبدأ المفحوصون الذين يعطون مستويات مختلفة وينتهون منها باستخدام مجموعات مختلفة من الأسئلة، أو بلغة البطارية «وحدات قياس» modules.

وتقيس الطبعة متعددة المستويات ثمانية مستويات يتألف كل منها من عشرة اختبارات مصنفة إلى ثلاث بطاريات هى:

١ – البطارية اللفظية وتتكون من اختبارات المفردات، وتكملة الجمل، والتصنيف اللفظى، والقياس التمثيلى اللفظى analogy.

٢ – البطارية الكمية وتتكون من اختبارات العلاقة الكمية وسلاسل الأعداد، وبناء المعادلات.

٣ – البطارية غير اللفظية وتتكون من اختبارات تصنيف الأشكال والقياس التمثيلى للأشكال، وتركيب الأشكال.

وقد قننت البطارية على عينة جيدة التمثيل للمجتمع الأمريكى فى المستويات الصفية العشرة المتضمنة فيها، وتتحول الدرجات الخام فيه إلى مقياس موحد فى جميع المستويات، والمعايير المستخدمة فى كل بطارية هى الدرجات المعيارية الاعتدالية متوسطها ١٠٠ وانحرافها المعيارى ١٦ (أى نسبة ذكاء انحرافية)، كما تستخدم المئينيات والتساعيات، ويمكن الحصول على مؤشر موحد من البطاريات الثلاث، وحسبت معاملات الثبات بطريقة كيودر- ريتشارد سدون وبلغت فى المتوسط بين ٠,٧ ، ٠,٨ ، أما معاملات الثبات بطريقة إعادة الاختبار فكانت أقل، وبصفة عامة فإن البطارية اللفظية أكثر ثباتا من البطاريتين الأخريين.

ومن حيث الصدق فإن معاملات الارتباط بين درجات البطارية عالية وكذلك معاملات الارتباط بين الاختبارات الفرعية، وتؤكد التحليلات العاملية للبطارية وجود عامل عام فى البطاريات الثلاث قد يدل على القدرة على الاستدلال باستخدام المحتوى المجرد والرمزى، وحسب صدق الاختبار أيضاً باستخدام محكات متعددة منها مقاييس المهارات الأساسية والاختبارات التحصيلية والكفاءة التعليمية، وكانت معاملات ارتباط البطارية اللفظية كالمعتاد أعلى من البطاريتين الأخريين، وعموما

التقويم النفسى ـــــــ

العقلى. وذلك بترتيب التلاميذ الذين اختبرهم من كل عمر على حسب درجاتهم فى الاختبار، ثم قسمهم إلى فئات متفاوتة فى المستوى العقلى هى أ. ب. جـ + جـ، جـ - د، هـ على الترتيب. والفئة (جـ) تدل على المتوسط بالطبع، ويبين فى الجدول حدود الدرجات لكل طبقة من طبقات الذكاء، ولما كان تلاميذ هذه الفئة من عمر واحد تقريبا فإنه كلما كان المستوى العقلى للفئة أعلى كان ذكاء أفرادها أكبر، وبالطبع لاتطابق «طبقات» الذكاء هذه المراتب التى تتحدد بنسب الذكاء وإن كانت تقترب منها كما هو مبين بالجدول رقم (٢٨).

جدول رقم (٢٨) يبين طبقات الذكاء ونسب الذكاء المقابلة لها

كما حددها القبانى (٨ : ٤٢)

الطبقة	نسب الذكاء المقابلة لها	
	من	إلى
هـ	٠٠	٧٨
د	٧٩	٨٨
جـ -	٨٩	٩٥
جـ	٩٦	١٠٤
جـ +	١٠٥	١١١
ب	١١٢	١٢١
أ	١٢٢	فما فوق

وقد حسب ثبات هذا الاختبار بطريقة إعادة الاختبار على مجموعات من التلاميذ تمثل أعمار ٩، ١٠، ١١، ١٢ فتراوحت المعاملات بين ٠,٨٦، ٠,٩٠ أما الصدق التلازمى للاختبار فقد تحددت بمعامل الارتباط بينه وبين اختبار جوادنف لرسم الرجل فوجد أن معاملات الارتباط تراوحت بين ٠,٤٧، ٠,٧٣ لعينات من أطفال أعمار ٩، ١٠، ١١. ويزيد مقدار هذا الارتباط كلما صغر عمر الأطفال، كما حسب معامل الارتباط بين درجات الاختبار ودرجات متاهة بورتيوس فبلغ ٠,٥١

لأطفال الصفوف الابتدائية، ٠,٤٧ لتلاميذ من عمرى ٩، ١٠ سنوات من فرق مختلفة، بالإضافة إلى دراسة علاقته بالنجاح المدرسى وتقديرات المدرسين للذكاء، وكانت العينات فى كل الأحوال صغيرة العدد.

والواقع أن هذا الاختبار من الاختبارات الجيدة حقا، وكان يمثل بداية موفقة لحركة القياس العقلى فى مصر، وقد ظهرت طبعته الثانية عام ١٩٣٩، ولو أنه توبع متابعة جادة كما حدث لمقياس بينيه الأصلى فى فرنسا. أو ستانفورد- بينيه فى أمريكا، أو بالارد فى إنجلترا لقدم علم النفس فى مصر خدمة جليلة، وفى رأينا أنه لو توافر عليه جهد جديد لإعادة تقنينه وتعديله بعد انقضاء ما يقرب من نصف قرن على تقنينه الأخير لكان فى ذلك إسهام كبير لحركة القياس العقلى فى وطننا.

٢ - **اختبارالذكاء المصور:** هذا الاختبار من إعداد أحمد زكى صالح، وجرب على مستويات عمرية مختلفة وتأكد أنه يصلح لقياس الذكاء ابتداء من سن الثامنة ولذلك أعدت معايير الاختبار (فى صورة مئينيات ونسب ذكاء تقليدية) من سن ٨ سنوات حتى ١٧ سنة، وهو اختبار متجانس الأسئلة بمعنى أنه يتألف من ٦٠ سؤالا كل منها يسأل المفحوص أن يبحث عن الشكل المخالف فى كل مجموعة من الأشكال تتكون من ٥ أشكال.

واستخرجت معاملات الثبات بطريقة التجزئة النصفية وتحليل التباين وتراوحت بين ٠,٧٥ و ٠,٨٥.

أما عن صدق الاختبار فقد ارتبط بالنتاج العقلى العام (كما يقاس باختبار القدرات العقلية الأولية) بمقدار ٠,٣٤ وتشبع فى عدد من البحوث العاملية التى أشرف عليها مؤلفه بالعامل العام تشبعات تتراوح بين ٠,٤٨ ، ٠,٦١ والاختبار فى رأى فؤاد أبو حطب (٢٨) يقيس إحدى قدرات الذكاء وهى التصنيف ويمكن أن يعد مقياسا لهذه القدرة وحدها ولا يتجاوز ذلك إلى الذكاء بمعناه الشامل، بالإضافة إلى أن محتواه يتأثر بالعوامل الثقافية، وبالرغم من هذه الحدود التى تقيد استخدام هذا الاختبار كمقياس للذكاء إلا أنه لايزال يستخدم فى كثير من البحوث الحديثة على أنه كذلك، كما لوحظ فى العديد من البحوث استخدامه لأعمار تفوق العمر الزمنى المحدد له، ناهيك عن قصور بيانات تقنينه التى لم تتجدد منذ أوائل السبعينات بعد وفاة مؤلفه.

٣- اختبار الذكاء غير اللفظى: هذا الإختبار من اعداد عطيه محمود هنا عن الأصل الأجنبى للإختبار العقلى المتعدد غير اللغوى لترمان ومككال ولورج Non-Language Multi - Mental Test

ويشبه هذا الاختبار الاختبار السابق (اختبار الذكاء المصور) فى أنه يتطلب فى جميع أسئلته - وعددها ٦٠ - أن يحدد الطفل الشكل المخالف من بين ٥ أشكال، ومعنى ذلك أنه متجانس التكوين ويقيس نفس عملية التصنيف التى أشرنا إليها بالنسبة للاختبار السابق، وحسب ثباته بمعادلة كيودر- ريتشارد سون لفئات أعمار من ٧,٥ سنة إلى ١٤,٥ سنة فتراوحت بين ٠,٧٢، ٠,٨٣ أما صدق الاختبار فقد تحدد بمعاملات ارتباطه باختبار الذكاء الثانوى (للقبانى) فبلغ ٠,٦٥ وقنن الاختبار على عينة من ٨٤١ طفلا فى المرحلتين الابتدائية والإعدادية تتراوح أعمارهم بين ٦ سنوات ٩ شهور و١٦ سنة ١١ شهرا.

وحسبت معايير الاختبار فى صورة أعمار عقلية يمكن منها حساب نسب الذكاء بالمعنى التقليدى وفى حالة الأطفال الذين يبلغون سن ١٤ سنة أو أكثر يكون العمر الزمنى فى معادلة نسب الذكاء مقدارا ثابتا هو ١٤ أو ١٦٨ شهرا، وهو إجراء تجاوزه المنطق السيكومترى الحديث، وينطبق على هذا الاختبار ما قلناه عن الاختبار السابق من حيث ضيق نطاق القدرات التى يقيسها من الذكاء وعدم تحديث بيانات تقنينه.

٤ - اختبار الذكاء الإعدادى: يهدف هذا الاختبار إلى قياس «القدرة على الحكم والاستنتاج خلال ثلاثة أنواع من المواقف: مواقف لفظية ومواقف عددية ومواقف تتناول الأشكال المرسومة» (٢٠ : ٤) وهو من إعداد السيد محمد خيرى الذى يرى أنه يقترب من مفهوم العامل العام عند سبيرمان، رغم أنه اختبار غير متجانس التكوين، وهو بذلك أقرب إلى اختبارات الذكاء من نمط ستانفورد - بينيه.

وقد بدأ تقنين هذا الاختبار عام ١٩٥٨ واشتملت عينة التقنين على ٣٤١٢ تلميذا من مدارس القاهرة و١٨٠٣ تلميذا من مدارس الوجه البحرى و٥٨٨ تلميذا من مدارس الوجه القبلى (أى أن العينة الكلية بلغت ٥٨٠٣ تلميذا) وحسب الثبات بطريقة إعادة الاختبار (ن = ٥١٤) فبلغ ٠,٩٢، وبطريقة التجزئة النصفية فبلغ ٠,٩١ أما فى تحديد الصدق فقد حسب معامل ارتباط الاختبار باختبار الذكاء الابتدائى (للقبانى) فوصل

إلى ٠,٦٥ بالإضافة إلى حساب معامل ارتباطه بتقديرات المدرسين للذكاء فبلغ ٠,٥٠٤ وحسبت معايير الاختبار فى صورة أعمار عقلية، وقد استخدم تقنين الاختبار فى عدد من رسائل الماجستير والدكتوراه إلا أنه فى حاجة إلى إعادة تقنين.

اختبارات ذكاء طلاب المرحلة الثانوية والجامعات:

من أشهر الاختبارات التى تقيس ذكاء طالب المرحلة الثانوية اختبار المستوى الثانى من اختبارات قدرات المرحلة الثانوية والجامعية (SCAT) School and College Ability Tests وله صورتان أ، ب. ويهدف هذا الاختبار إلى التنبؤ بالتحصيل الأكاديمى، ولذلك يعطى للمفحوص فيه ثلاث درجات، درجة فى الاختبارات اللفظية، ودرجة فى الاختبارات الكمية، ودرجة كلية، وتعتمد الدرجة اللفظية على اختبارين هما: فهم الجمل ومعانى الكلمات، وتعتمد الدرجة العددية على اختبارين أيضاً هما: العمليات العددية والمسائل العددية (حل المشكلات العددية).

ويتفق هذا الاختبار مع النظرية الحديثة لقياس ما يسمى «القدرات المستثمرة» (٢٨ : ٣٤)، ولذلك فهو يعتمد على المعرفة اللغوية والعمليات الحسابية المتعلمة فى المراحل التعليمية السابقة، وهو فى هذا لايختلف عن كثير من اختبارات الذكاء الأخرى التى تعكس مقادير متفاوتة من آثار «التعلم» ولاتقيس الوسع الفطرى» مستقلا عن الخبرات السابقة.

ويمكن لدرجات الاختبارات اللفظية والكمية والكلية أن تصاغ فى ضوء مقياس مشترك يسمح بالمقارنة بينها من مستوى لآخر من المستويات التى يقيسها الاختبار، كما يمكن لهذه الدرجات أن تتحول إلى مئينيات أو تساعيات مشتقة من عينة تقنين شاملة لكل مستوى فيه.

وقد حسب ثبات الاختبار بطريقة كيودر ـ ريتشاردسون وتراوحت معاملات الثبات بين ٠,٨٨ ، ٠,٩٦ للمقاييس الثلاثة، وقد أجرى على هذا الاختبار تحليل دقيق للمفردات للحصول على تمييز واضح بين الفئتين اللفظية والكمية، وتؤكد النتائج تمايز الاختبار تمايزا متصاعدا مع تزايد العمر والمستوى التعليمى بما يتفق مع نظريات التنظيم العقلى، أما عن صدق الاختبار فقد اهتم مؤلفوه بصدقه التنبؤى واستخدموا لهذا الغرض محكات التحصيل المدرسى، وكانت معاملات الارتباط مع الاختبارات المدرسية أعلى منها مع

التقديرات ، وكانت فى المستويات الدنيا أعلى منها فى المستويات العليا.

وقد أعدت مجموعة من الاختبارات العقلية لتستخدم فى أغراض انتقاء طلاب الجامعات وتصنيفهم وإرشادهم وتوجيههم، ومن أشهرها اختبار الاستعداد التعليمى Scholastic Aptitude Test (SAT) من إعداد هيئة امتحانات القبول بالجامعات فى الولايات المتحدة College Entrance examination Board (CEEB) والذى نشر لأول مرة عام ١٩٢٦ . وتظهر صور عديدة جديدة من هذا الاختبار كل عام، ويحصل المفحوص على درجات منفصلة للقسم اللفظى والقسم الرياضى فى الاختبار، ويعطى هذا الاختبار تقديرا مبدئيا لاستعداد طالب المرحلة الثانوية للعمل الجامعى.

ومن هذه الاختبارات أيضاً امتحان برنامج القياس الجامعى American Col-lege Testing Program (CAT) ويتضمن ٤ أقسام هى: استخدام اللغة، استخدام الرياضيات، قراءة المواد الاجتماعية، قراءة العلوم الطبيعية، وقد أعد هذا الاختبار لندكويست Lindquist ويعكس وجهة نظره فى أن الامتحان يجب أن يكون عينة من العمل الجامعى، وتتداخل فيه اختبارات الاستعدادات التقليدية واختبارات التحصيل ويركز على المهارات العقلية الأساسية المطلوبة للأداء الناجح فى الجامعات.

ويوجد اختبار ثالث هو الاختبارالسيكولوجى لجامعة أوهايو: Ohio (OSUPT) State University Psychological Test وهو اختبار قوة محض، يطبق دون زمن محدد، ويستغرق فى العادة بين ١,٥ ساعة وساعتين، كما أنه فى جوهره اختبار لفظى يتضمن الأضداد والمتماثلات اللفظية وفهم القراءة والعلوم الطبيعية والرياضيات والعلوم الاجتماعية والأدب.

ومن الاختبارات الحديثة فى الظهور اختبار التأهيل الجامعى Collega Qual-ification Test . ويعطى درجات فى القسم اللفظى والعددى والمعلومات العلمية ومعلومات العلوم الاجتماعية، والمعلومات الكلية، والدرجة الكلية فى الاختبار، وكذلك برنامج الاختبارات للكليات الأمريكية American College Testing Pro-gram (ACT) والذى يشمل اختبارات الاختيار الأكاديمى ويتألف من أربعة أقسام: الاستخدام اللغوى، والاستخدام الرياضى، قراءة الدراسات الاجتماعية، قراءة العلوم الطبيعية.

ومن الواضح أنه باستثناء اختبار SAT (الذى تكمله الاختبارات التحصيلية) فإن

جميع هذه الاختبارات تمثل مزيجا من الاستعدادات الهامة ومعرفة مادة دراسية من المجالات الأكاديمية الأساسية، وحين تستخدم الدرجات المنفصلة فإن صدقها الفارق فى التنبؤ بالتحصيل فى الميادين المختلفة يصبح موضع شك، ومعنى ذلك أن الدرجة الكلية هى أفضل منبئ بالأداء فى جميع المقررات الجامعية تقريباً، وإذا كان لابد من استخدام الدرجات الجزئية فإن الدرجات اللفظية هى أفضل المنبئات الفردية، ويجب أن ننبه إلى أن هذه الاختبارات لا تستخدم كبدائل لاختبارات نهاية المرحلة الثانوية (الثانوية العامة فى بلادنا) فى التنبؤ بالتحصيل الجامعى، فالواقع أن الدراسات الحديثة (٦١ : ٢٣٢) تؤكد أن درجات المرحلة الثانوية يمكن أن تتنبأ بالتحصيل الجامعى بنفس كفاءة هذه الاختبارات، إن لم يكن أفضل، وحين ترتبط درجات اختبارات الذكاء هذه بدرجات نهاية المرحلة الثانوية (الثانوية العامة) فإن التنبؤ بالأداء الجامعى يصبح أكثر دقة.

ومن الاختبارات التى أعدت فى مصر لقياس ذكاء تلاميذ المرحلة الثانوية وطلاب الجامعات ما يأتى:

(١) **اختبار الذكاء الثانوى**: وقد أعده إسماعيل القبانى ويتألف من ٥٨ سؤالا تتناول سلاسل الأعداد وإنتاج الجمل وإدراك العلاقات بين الكلمات وإدراك السخافات والاستدلال، ويصلح للتطبيق على تلاميذ المرحلتين الاعدادية والثانوية فى المدى العمرى من ١٢ سنة إلى ١٨ سنة وعلى الكبار ممن تزيد أعمارهم على ١٨ سنة، وقد أعدت معايير الاختبار فى صورة طبقات للذكاء فى المستوى (أ) وهو أعلى المستويات ، ثم المستوى (ب، جـ، د) ثم المستوى (هـ) وهو أدنى المستويات، كما أعد جدول يتضمن الدرجات التى تحدد كل طبقة من كل مستوى عمرى، ولاتتوافر لدينا بيانات سيكومترية عن هذا الاختبار.

(٢) **اختبار الذكاء العالى**: من إعداد السيد محمد خيرى على أساس مفهوم سبيرمان للعامل العام على أنه إدراك العلاقات والمتعلقات ويتشابه مع اختبار الذكاء الإعدادى فى أنه يقيس القدرة على الحكم والاستنتاج خلال ثلاثة أنواع من المواقف : لفظية وعددية ومواقف تتناول الأشكال المرسومة، وقد قنن على عينة عددها ٥٨٢٨ طالبا فى المدارس الثانوية والمعاهد العليا والجامعات ، واشتملت العينة على المستوى الثانوى بفرعيه العلمى والأدبى، كما اشتملت عينة المستوى الجامعى على كليات نظرية وأخرى عملية.

وحسب ثبات الاختبار بطريقتى إعادة الاختبار والتجزئة النصفية فبلغ معامل الثبات ٠,٨٤٥ ٠,٨٨١ على التوالى، أما فى تقدير الصدق فقد حسب معامل ارتباطه باختبار الذكاء الثانوى فبلغ ٠,٦٩٤ كما حسب ارتباطه بتقديرات المدرسين لذكاء تلاميذهم فى التعليم الثانوى فبلغ ٠,٥٢٢ وأخيراً حسب ارتباطه بدرجات أفراد العينة الذين فى السنة الإعدادية أو الأولى من المستوى العالى أوالجامعى فى نهاية العام فكان ٠,٥١٨ وأعدت للاختبار معايير فى صورة مئينيات ، وهذا الاختبار حالة كحال سابقه اختبار الذكاء الإعدادى يحتاج إلى إعادة تقنين.

(٣) اختبارالاستعداد العقلى للمرحلة الثانوية والجامعات:

(هذا الاختبار من إعداد رمزية الغريب وينقسم إلى ٥ أقسام هى):

اليقظة الذهنية، التفكير الرياضى، إدراك العلاقات المكانية، فهم الرموز اللغوية، والتفكير المنطقى، وهو من اختبارات القوة، ويحتاج فى المتوسط إلى ساعة ونصف للانتهاء منه، وقد حسب الثبات بطريقة التجزئة النصفية (ن = ٥٥٠) فبلغ معامل الارتباط ٠,٩٢ واستخدمت طريقتان لحساب الصدق التلازمى هما حساب معامل ارتباطه باختبار القدرات العقلية الأولية فبلغ ٠,٧٧ والمقارنة بين المجموعات المتضادة وكانت الفروق بين الممتازين عقليا، والمتخلفين عقليا لها دلالة إحصائية.

اختبارات طلاب الدراسات العليا:

يمكن القول أن الاختبارات التى تستخدم فى قياس طلاب الدراسات العليا تمثل مزيجا من اختبارات الذكاء العام واختبارات التحصيل، ومن أشهر هذه الاختبارات امتحانات سجل الدراسات العليا Graduate Record Examinations (GRE) والتى تطبق على الطلاب المتقدمين للالتحاق بالدراسات العليا فى أمريكا، وقد ظهرت هذه السلسلة من الاختبارات عام ١٩٣٦ بهدف توجيه الطلاب وقياس كفاءة التدريس وانتقاء أفضل العناصر للحصول على المنح والبعثات والتعيين فى وظيفة معيد.

ويتكون الاختبار فى إحدى صوره من اختبار للاستعداد واختبار تحصيلى فى ميدان تخصص الطالب من مستوى رفيع مثل الأحياء والسياسة والفلسفة وعلم النفس واللغة الإسبانية...إلخ،. والاختبار فى جوهره اختبار استعداد دراسى أو ذكاء عام ملائم لطلاب الدراسات العليا فى بداية التحاقهم بها.

وللاختبار صورة أخرى لا تتألف فقط من اختبارات استعدادات وتحصيل رفيع، وإنما تشمل كذلك ثلاثة اختبارات من نوع اختبارات «المجال» هى العلوم الاجتماعية والإنسانيات والعلوم الطبيعية، وهذه الاختبارات تؤكد فهم المفاهيم الأساسية والقدرة على تطبيق المعلومات وليس مجرد استدعاء الحقائق التفصيلية.

وتحول درجات هذا الاختبار إلى درجة معيارية معدلة متوسطها ٥٠ وانحرافها المعيارى ١٠. وحسب له الثبات بطريقة كيودر- ريتشاردسون وطريقة التجزئة النصفية، وتتراوح معاملات الثبات بين ٠,٨٤ ، ٠,٩٥ أما الثبات بطريقة إعادة الاختبار فيتراوح بين ٠,٨٥ ، ٠,٩٧ بفاصل زمنى مقداره عام واحد، وبين ٠,٧١ ، ٠,٩٣ بفاصل زمنى مقداره عامان، وحسب الصدق التنبؤى فى ضوء محكات معينة مثل الدرجات الأكاديمية فى الدراسة العليا، والنجاح فى مقابل الفشل فى الدراسة، وتقديرات الأساتذة ، وأداء الطلاب فى امتحانات التأهيل للدكتوراه.

اختبارات ذكاء الراشدين المتفوقين:

يمكن القول أن اختبارات طلاب الجامعة والخريجين وطلاب الدراسات العليا تلائم مستوى الراشدين المتفوقين لأغراض التقويم والبحث فى الميدان المهنى، ومن ذلك أن اختبار ميللر للقياس التمثيلى Miller Analogies وهو يصلح لتقويم طلاب الدراسات العليا. يمكن أن يستخدم فى هذا الغرض، ويتألف من متماثلات مركبة مصاغة فى محتوى أكاديمى وعلى درجة كبيرة من الصعوبة، مما يجعله ينتمى إلى فئة اختبارات القوة، ومن هذه الاختبارات أيضاً اختبار اتقان المفاهيم Concept Mastery Test وقد ظهر نتيجة البحوث الطولية الرائدة التى قام بها ترمان فى ميدان المتفوقين، ويتألف من أسئلة القياس التمثيلى والمتشابهات والأضداد التى تنتمى إلى محتوى متنوع فى مجالات العلوم الفيزيائية والبيولوجية والرياضيات والتاريخ والموسيقى وغيرها.

اختبارات الذكاء المتحررة من أثر الثقافة:

تسعى هذه الإختبارات الى التحكم فى بعض الأبعاد التى تتألف منها الثقافة Culture، والتى تختلف فيها الثقافات. وعلى وجه الخصوص بعد اللغة، فإذا كان المفحوصون يتحدثون لغات مختلفة يجب أن يتحرر الاختبار من هذا العنصر سواء من

جانب الفاحص أو المفحوص، أما حين تختلف الأطر التعليمية اختلافات واسعة وخاصة فى المجتمعات التى تنتشر فيها الأمية، كأغلب الدول النامية ـ فإن الاختبارات يجب أن تتحرر من عنصر القراءة ـ كما أشرنا إلى ذلك من قبل. وفى حالة الاختبارات المتحررة فى أثر القراءة Non-Reading لايستبعد أثر اللغة «الشفوية» بالطبع لأن هذه الاختبارات تعد لجماعات تتحدث بنفس اللغة، ومن العوامل الأخرى التى تختلف فيها الثقافات عامل السرعة speed، وتتفاوت فى ذلك ابتداء من الإيقاع الشخصى Personal tempo وحتى الدافعية للعمل بسرعة وتقدير العمل السريع، وفى هذا تختلف الثقافات الفرعية للمجتمع الواحد (وخاصة بين الريف والمدينة)، وأخيرا فإن الثقافات المختلفة (فرعية أو عامة) تختلف فى التعامل مع محتوى الاختبارات، وقد حاول بعض المؤلفين إعداد اختبارات غير لغوية وغير قرائية، ولكنها تتألف من صور لأشياء واقعية كالأثاث والألعاب والآلات (راجع فى ذلك اختبارات المقياس العملى لوكسلر، واختبارات الذكاء المصور والذكاء غير اللفظى) وبالطبع فإن هذا المحتوى «السيمانتى» للصور يعتمد على إطار الخبرة السابقة.

ولهذه الأسباب وغيرها ـ ظهرت اختبارات الذكاء المتحررة من أثر الثقافة Cul-ture-free أو الاختبارات التى تصلح لأغراض المقارنة بين الثقافات Cross cultural وكان أول اختبار ظهر فى الميدان اختبار بيتا الحربى أثناء الحرب العالمية الأولى، وفيه تم التحكم فى العنصر اللغوى سواء من جانب الفاحص أوالمفحوص، وكانت تعطى التعليمات بالإشارة والإيماءات والتمثيل، أما استجابات المفحوص فكانت تتمثل فى مجرد وضع علامات بسيطة (فيما عدا اختبار واحد كان يطلب من المفحوص كتابة الأعداد). وقد أعيد تقنين هذا الاختبار ـ للأغراض غير العسكرية ـ عام ١٩٤٦، وأصبح يتألف من ٦ اختبارات فرعية هى :المتاهات، وتعويض الرموز والأعداد، والسخافات المصورة، ولوحة الأشكال المصورة، وتكميل الصور، والسرعة الإدراكية، ويتطلب فى هذا التعديل بعض التعليمات اللفظية إلا أنه يعتمد اعتمادا كبيرا على الأسئلة التدريبية.

وقد ظهرت عدة اختبارات فى هذا الميدان منها اختبارات سبيرمان الحسية للذكاء التى أعدها باللغة العربية إسماعيل القبانى وعبدالعزيز القوصى، والمقياس الأدائى العالمى لليتر Leiter International Performance وسوف نقتصر على ٣ اختبارات مهمة هى: اختبارات كاتل المتحررة من أثر الثقافة، واختبار المصفوفات المتتابعة

لرافن واختبار رسم الرجل لجود انف.

(١) اختبارات كاتل غير المتحيزة للثقافة: أعد هذه الاختبارات العالم الأمريكى المعاصر ريموند كاتل، ولها ثلاثة مستويات: المقياس الأول للأعمار من ٤ سنوات إلى ٨ سنوات والراشدين المتخلفين عقليا، والمقياس الثانى للأعمار من ٨ سنوات إلى ١٢ سنة والراشدين العاديين، والمقياس الثالث من سن ١٣ سنة إلى ١٩ سنة والراشدين المتفوقين، ولكل مقياس من هذه المقاييس الثلاث صورتان متكافئتان. ويتطلب المقياس الأول تطبيقا فرديا أو جماعيا، ويتألف المقياس الأول من ٨ اختبارات، أربعة منها فقط تعد من نوع الاختبارات المتحررة ثقافيا، أما الاختبارات الأربعة الأخرى فتتطلب كلا من الفهم اللفظى والمعلومات الثقافية الخاصة، أما المقياسان الثانى والثالث فمتشابهان فى المحتوى ومختلفان فى مستوى الصعوبة، ويتألف كل منهما من أربعة اختبارات هى: سلاسل الأشكال، وتصنيف الأشكال، ومصفوفات الأشكال، والشروط، وتتوافر للمقياس الأول فقط معايير نسب الذكاء الانحرافية بانحراف معيارى ٢٤ أو ١٦، وقد شاع التحويل الأخير حتى يسمح للمقاييس بالمقارنة بغيرها.

وقد نقل المقياس الثانى إلى اللغة العربية أحمد عبد العزيز سلامة وعبد السلام عبدالغفار، ولا تتوافر عنه - فى حدود علمنا - بيانات تقنين، أما المقياسان الثانى والثالث فقام بإعداده فؤاد أبو حطب وآمال أحمد مختار صادق كما قام بتقنينه عدد من طلاب الدراسات العليا فى مصر فى بحوثهم للماجستير والدكتوراه، كما استخدمت آمال صادق المقياسين الثانى والثالث فى دراستين عامليتين للقدرة الموسيقية (٩) وأكدت النتائج تشبع المقياسين بالعامل العام فى عينتين مصريتين إحداها من الأطفال من مرحلة التعليم الاساسى (ن = ١٨٥) والأخرى من المراهقين والراشدين (ن = ٢٣١)، وقياس هذه الاختبارات للعامل العام هو الغرض الأساسى الذى تقوم عليه اختبارات كاتل متأثرة فى ذلك بنظرية أستاذه سبيرمان.

كما قامت نادية السيد الحسينى فى رسالتها للدكتوراه تحت اشراف آمال صادق عام ١٩٨٧ بإعادة تقنين المقياسين الثانى والثالث على عينات من المرحلة الابتدائية (ن= ٣٠٠، والإعدادية (٢٠٠) طالبا، والثانوية(ن = ٨٠٠ طالبا)، أما التقنين الحالى

* شارك في التقنين الثالث ٢٠٠٥ أ.د. مصطفى عبد العزيز

* والذى أجرى عام ٢٠٠٦ (ن = ٦٤٢ من المراحل الإعدادية والثانوية والجامعية) ونشرعن كل هذه الإجراءات فى دليل وكراسة التعليمات التى اعدتها آمال صادق عام ٢٠٠٦ .

وتبين كراسة تعليمات الإختبار الخصائص السيكومترية للإختبارين الثانى والثالث حيث تم حساب الثبات بطرق متعددة فى دراسات التقنين الثلاثة فقد تراوحت قيمة معامل الثبات بطريقة الصور المتكافئة بين ٠,٥٦ ، ٠,٨٧ ، وبطريقة التجزئة النصفية بين ٠,٧٢ ، ٠,٨٤ أما الثبات العاملى لاشتراكيات العوامل للدرجة الكلية فى دراسة آمال صادق عام ٦٨ فبلغت ٠,٦٣ ، ٠,٩٥ على التوالى فى دراسة الأطفال والمراهقين والراشدين.

أما صدق الإختبارات فقد تم تحديده باستخدام التحليل العاملى (آمال صادق ٦٨) وكانت تشبعات الاختبارات الفرعية دالة وكذلك الدرجة الكلية حيث تراوحت بين ٠,٥٧ ، ٠,٩٣ . كما تم تحديد الصدق المرتبط بالمحكات (اختبارات أخرى للذكاء، التحصيل الدراسى- ومحك تمايز العمر) وكانت معاملات الارتباط متباينة فى الدراسات الثلاث للتقنين.

أما معايير الإختبارات (المقياس الثانى والثالث) فقد تم تضمين كراسة التعليمات المتوسطات والإنحرافات المعيارية للعينات المختلفة حتى يمكن حساب الدرجات المعيارية ونسبة الذكاء الإنحرافية بمتوسط مقداره ١٠٠ وانحراف معيارى ١٦ حتى يمكن المقارنة بين اختبارات الذكاء المختلفة.

وفي دراسة أخرى قام بها هشام فتحي جاد الرب* موضوعها: تطوير اختبار كاتل للذكاء باستخدام نماذج السمات الكامنة وأثر ذلك على قدرة الاختبار على التنبؤ بالتحصيل الدراسى وذلك للحصول على درجة الماچستير، وتحدد الهدف من الدراسة: تدريج مفردات اختبار كاتل للذكاء – المقياس الثانى باستخدام نموذج راش، والتحقق من مدى موضوعية التفسير لنتائج الاختبار بعد تدريجه، وكذلك دراسة أثر التدريج على رفع القيمة التنبؤية للاختبار باستخدام محك التحصيل الدراسى.

وللتحقق من صحة فروض الدراسة، استخدم الباحث عينة مكونة من ١٠١٥

* أشرف علي الرسالة أ.د. ممدوح عبد المنعم الكنانى.

طالبا وطالبة من طلاب المرحلة الإعدادية (٥٣٧ طالبة، ٤٧٨ طالبا) من الصف الثانى الإعدادى، وقد قام الباحث بإعداد اختبار تحصيلى فى مادة العلوم فى الفصل الدراسى الأول. وتم تحليل مفردات اختبار كاتل للذكاء وتدريجها باستخدام برنامجى مايكروسكيل Microscale وراسكال Rascal بالإضافة لتحليل النتائج باستخدام برنامجى SYSTAT, SPSS وتم التحقق من توافر افتراضات نموذج راش فى البيانات وتم حذف الأفراد والمفردات غير الملائمة.

وقد توصلت الدراسة لعدم تأثر تقديرات صعوبة مفردات الاختبار بعد تدريجها بتغير قدرة الأفراد المستجيبين عليها، وكذلك عدم تأثر تقديرات قدرة الأفراد بمجموعة المفردات التى يستجيبون عليها، وقد أنعكس ذلك على ارتفاع قدرة الاختبار على التنبؤ بالتحصيل الدراسى. كما توصلت الدراسة لمعايير جديدة للاختبار لتحويل الدرجة الخام على الاختبار إلى تقدير للقدرة بوحدة اللوجيت.

(٢) اختبارات المصفوفات المتتابعة لرافن: ظهر هذا الاختبار فى إنجلترا عام ١٩٣٨ لقياس العامل العام عند سبيرمان أيضًا، وتتطلب فى جوهرها إدراك العلاقات بين وحدات مجردة، وهو فى رأى معظم علماء النفس فى بريطانيا أفضل مقاييس العامل العام ولهذه الاختبارات ثلاث مستويات ظهرت فى ثلاثة اصدارات المصفوفات المتتابعة الملونة The coloured Progressive Matrices Test، المصفوفات المتتابعة العادية The Standard Progressive Matrices Test، والمصفوفات المتتابعة المتقدمة The Advanced Progressive Matrices Test.

اختبار المصفوفات المتتابعة الملون:

ويتألف المقياس الأول من ٣٦ مفردة مقسمة إلى ثلاث مجموعات أ، أب، ب، والمجموعة الأولى والثالثة هى ذاتها المجموعة الأولى والثانية من المقياس الثانى، ويصلح هذا المقياس للأطفال من سن ٥ سنوات الى ١١ سنة، والألوان المستخدمة فقط لجذب انتباه الأطفال. وفى كل مفردة تعرض على المفحوص شكل به جزء ناقص وعلى المفحوص تحديد هذا الجزء الناقص باختياره شكلا من مجموعة أشكال مطروحة أسفل الشكل الأساسى.

اختبار المصفوفات المتتابعة العادى:

ويتألف الاختبارا من ٦٠ مصفوفة (أو رسما) حذف منها أحد أجزائها، وعلى

المفحوص أن يحدد المحذوف بالانتقاء من بين ٦ أو ٨ بدائل، وتتجمع الأسئلة فى ٥ سلاسل، يحتوى كل منها على ١٢ مصفوفة متزايدة الصعوبة ولكنها متشابهة من حيث المبدأ، وتتطلب المصفوفات الأولى دقة فى التمييز، أما المصفوفات المتأخرة والأكثر صعوبة فتتضمن إدراك العلاقات المنطقية .كما تتطلب كل مجموعة من المجموعات الخمس نمطا مختلفا من الإستجابة:

فالمجموعة (أ) تتطلب تكملة نمط أو مساحة كاملة.

والمجموعة (ب) تتطلب تكملة نوع من قياس التماثل بين الأشكال.

والمجموعة (ج) تتطلب التغيير المنتظم في أنماط الاشكال.

والمجموعة (د) تتطلب اعادة ترتيب الشكل أو تعديله أو تغييره بطريقة منتظمة.

والمجموعة (هـ) تتطلب تحليل الأشكال الى اجزاء على نحو منتظم وادراك العلاقات بينها.

وهكذا ترتب المجموعات بطريقة متتابعة، وفى رأى راڤن أن هذا التتابع يتم حسب مستويات صعوبة أو تعقد العمليات العقلية المعرفية، فبينما تتطلب المجموعات الأولى والأكثر سهولة الدقة فى المقارنة والتمييز والتماثل، تتطلب المجموعات الأخيرة الأكثر صعوبة القدرة على ادراك العلاقات المنطقية. ويذكر راڤن فى هذا الصدد ان اختبار المصفوفات المتتابعة العادى يمكن أن يصنف بأنه اختبار للملاحظة والتفكير الواضح كما يهدف إلى شمول المدى الكلى للنمو العقلى ابتداء من المستوى الذى يستطيع الطفل عنده ادراك فكرة التعرف على الجزء الناقص الذى يكمل به النمط الشكلى ويمتد إلى قياس القدرة القصوى للشخص على المقارنة والإستدلال. وعلى الرغم من أن درجات الراشدين تميل إلى التجمع فى النصف العلوى من مجموعات المفردات فإنه يوجد عدد كاف من المفردات للتمييز بينهم (راڤن: ١٩٦٠ :٢).

ويعطى اختبار المصفوفات المتتابعة العادى للمفحوص بصرف النظر عن عمره الزمنى بنفس التتابع للمجموعات الخمس، ثم يطلب منه القيام بالعمل تبعا لمعدله الخاص ووفقا لسرعته الخاصة وبدون مقاطعة منذ أن يبدأ العمل وحتى الانتهاء منه، ويمكن أن يطبق فرديا أو جماعيا، وليس لهذا الاختبار زمن محدد (أى أنه ينتمى إلى اختبارات القوة).

اختبار المصفوفات المتتابعة المتقدم:

أعد هذا الإختبار للإستخدام مع الأفراد الذين تزيد أعمارهم عن ١١ سنة من ذوى القدرة العقلية العادية أو الممتازة، ويتألف من مجموعتين I، II . وتتألف المجموعة I من ١٢ مفردة معدة للتدريب على طريقة العمل، أما المجموعة II فتتألف من المجموعات جـ، د، هـ من الإختبار العادى. وتفيد المجموعة I فى تحديد ما إذا كان المفحوص متخلفا أو متوسطا أو ممتازا من الوجهة العقلية وذلك فى دقائق قليلة فإذا ظهر أن الشخص من ذوى القدرة العقلية المتوسطة أو الأعلى من المتوسط يمكن تطبيق المجموعة IIعليه حتى يكون التقدير أكثر دقة وكما ذكرنا أن هذه الاختبارات الثلاثة غير موقوتة إلا أن الزمن اللازم للإجابة على كل اختبار من الاختبارات الثلاثة من ١٥–٣٠ دقيقة (رافن ١٩٦٠ : ١) وان المفردة الأولى فى كل مجموعة من مجموعات الاختبارات بمثابة تدريب عن المجموعة المنتسبة لها.

وقد نقلت هذا الاختبار إلى اللغة العربية الهيئة الفنية للخدمة السيكولوجية العسكرية، ولا تتوافر لدينا بيانات عن خصائصه السيكومترية للعسكريين كما كتب الدارس العراقى قيس عبد الفتاح مهدى إلى أحد مؤلفى هذا الكتاب (فؤاد أبو حطب) حول الدراسات التى أجريت حول هذا الاختبار بمركز البحوث التربوية والنفسية بجامعة بغداد وخاصة المصفوفات الملونة الذى تجرى محاولة تقنينه على المرحلة الابتدائية.

وقد تم تقنين إختبار المصفوفات المتتابعة العادى على البيئة السعودية عام ١٩٧٧ بواسطة فريق من الباحثين بإشراف فؤاد أبو حطب (٣٢) حيث طبق الاختبار على عينة تتألف من ٤٩٣٢ مفحوصًا من البنين والبنات تمتد أعمارهم من ٨ سنوات وحتى ما بعد سن الثلاثين اختيروا من تلاميذ المرحلة الإبتدائية والمتوسطة وطلاب المرحلة الثانوية والمرحلة الجامعية والخريجين بالمنطقة الغربية بالمملكة العربية السعودية، وحسب صدق التكوين الفرضى للاختبار، حيث أكد صدق المحتوى أنه اختبار غير لغوى، كما أكدت معاملات الارتباط بين درجات الاختبار والزمن المستغرق فى أدائه أنه اختبار قوة (حيث لم تكن هذه المعاملات دالة). وثبت أيضًا صحة الترتيب التتابعى للمجموعات الخمس لأسئلة الاختبار باعتبارها متدرجة فى

الصعوبة وذلك من خلال المقارنة بين متوسطات عينات الأعمار المختلفة فى كل مجموعة من المجموعات الخمس للأسئلة بالنسبة لكل درجة كلية يحصل عليها المفحوص، ثم تأكد أيضًا أن تتابع المفردات داخل كل مجموعة من المجموعات الخمس منتظم من السهل إلى الصعب وذلك من خلال فحص معاملات سهولة المفردات فى الأعمار المختلفة، كما استخدم فى حساب صدق التكوين الفرضى طريقة تمايز العمر وتمايز الصفوف الدراسية ، وقدر كذلك الصدق المرتبط بالمحكات باستخدام بعض الاختبارات العقلية الأخرى والتحصيل الدراسى وكانت معاملات الارتباط دالة.

وحسب الثبات بطريقة كيودر- ريتشاردسون وطريقة إعادة الاختبار وكانت المعاملات مرتفعة ودالة.

وفى الدراسات التى أجريت على ثبات الاختبار تراوحت معاملات الثبات بطريقة التجزئة النصفية من ٠,٦٥ إلى ٠,٩٤ للاختبار الملون أما اعادة التطبيق فقد تراوحت المعاملات بين ٠,٧١ ، ٠,٩٣ (ساتلر ١٩٩٢: ٣٠٨) . كما أشار ساتلر أيضا ان الدراسات العاملية تشير إلى وجود العامل العام، كما أشارت بعض الدراسات إلى وجود أكثر من عامل وبصفة خاصة للإختبار الملون. أما معايير الاختبار فقد تم تحديد مئينيات له .

وقد أجريت فى مصر دراستان الأولى قامت بها آمال جلال محمد للحصول على درجة الماچستير فى علم النفس التربوى عام ٢٠٠٠ * هدفت فيها إلى تقنين اختبار المصفوفات المتتابعة الملون. وأجرت البحث على عينة قوامها ٣٤٣٣ تلميذا من مرحلة التعليم الاساسى يتراوح عمرهم من ٦-١١ سنة (ذكور ١٧٢٣ - اناث ١٧١٠) . وقد تحققت الباحثة من بنية المقياس وتتابع الاسئلة فى كل مجموعة من السهل إلى الصعب، كما قامت بتحديد صدق المقياس وثباته بطرق مختلفة. فاستخدمت لتحديد صدق المقياس محك تمايز العمر ومحك الاتساق الداخلى، والصدق المرتبط بالمحكات (اختبار رسم الرجل للعينات فى الأعمار المتتالية وكانت هذه المعاملات دالة فيما عدا فى العمر الزمنى ست سنوات ، تسع سنوات) كما كانت جميع المعاملات دالة عند مستوى ا. و بين درجات الاختبار والتحصيل الدراسى، وايضا تقديرات المعلمين.

* أشرف أ.د. على السيد خضر ،د. هشام محمد مخيمر.

أما ثبات الاختبار فتم تحديده باستخدام طريقة اعادة التطبيق فى العينات المختلفة العمر وكانت جميع المعاملات دالة عند مستوى٠,٠١ ، وكذلك عند استخدام التجزئة النصفية.

وقد درست الباحثة العلاقة بين الأداء الإختبارى والمستوى الاجتماعى والثقافى للأسرة والترتيب الميلادى للمفحوص.

أما معايير الاختبار فقد قامت الباحثة بحساب الدرجة التائية، والمئينيات وكان من الأفضل تحديد نسب الذكاء الإنحرافية وهى الدرجة المعيارية المعدلة التى تسمح بالمقارنة فى الأداء بين اختبارات الذكاء المختلفة والتى أصبحت أكثر شيوعا فى أغراض المقارنة.

وفى دراسة أخرى قام بها أمين صبرى نور الدين بعنوان «فاعلية استخدام الاختبار الموائم باستخدام الحاسب فى تقدير قدرة الأفراد وتحديد الخصائص السيكومترية للمقياس» للحصول على درجة الدكتوراه. وقد استخدم فى هذه الدراسة اختبارات المصفوفات المتتابعة بمستوياتها الثلاثة (الملون – العادى– المتقدم). وفيما يلى عرض للدراسة كما لخصها الباحث ذاته* .

بحثت دراسة أمين صبرى نور الدين مدى فعالية الاختبار التواؤمى المحوسب Computerized Adaptive Testingفى تقدير قدرة الأفراد Ability Estimation وتحديد الخصائص السيكومترية للقياس مقارنة بالاختبار الخطى المحوسب linear testing وذلك من خلال أداء المفحوصين على أحد اختبارات القدرة العقلية العامة. قام الباحث بتطبيق اختبار المصفوفات المتدرجة لرافن بإصداراته المختلفة الملونة والعادية والمتقدمة Raven, 1958, Raven, 1965, Raven, Court & Raven, 1983) بصورة ورقية على أكثر من (٥٠٠٠) طالب وطالبة بهدف حساب مَعْلَمات الاختبار Parameters، وبعد التأكد من أحادية البعد، ومطابقة البيانات للنموذج ثلاثى المعْلَم Three Parameter Logistic Model لكل من الاختبارين، تم حذف بعض المفردات، وإنشاء مستودع للمفردات Item Pool مؤلف من (١٣١) مفردة ذات قيم مَعْلَمات صالحة. ثم قام الباحث بإنشاء وتطبيق مجموعة من الاختبارات التواؤمية

* أشراف أ.د. فؤاد أبو حطب، د. بديوى إبراهيم علام، د. وفاء عبد الجليل خليفه.

المحوسبة بناء على مستودع المفردات حيث أجرى الباحث ثلاثة دراسات مستقلة على أكثر من (٥٠٠) طالب وطالبة آخرين طبق على كل منهم اختبارين محوسبين إحداهما تواؤميا والآخر خطيا. قارنت الدراسة الأولى بين فعالية الاختبار التواؤمى المحوسب الذى يتوقف بعد درجة محددة من الدقة مقارنة بالاختبار الخطى المحوسب. والدراسة الثانية قارنت بين فعالية الاختبار التواؤمى المحوسب الذى يتوقف بعد عدد محدد من المفردات مقارنة بالاختبار الخطى المحوسب. أما الدراسة الثالثة فقد قارنت بين اختبارين تواؤميين محوسبين يستخدم أحدهما تقدير القدرة بالأرجحية القصوى Maximum Likelihood Estimation والثانى بالبعدى الأقصى Maximum A Posteriori وكل منهما ينتهى بعد تطبيق عدد محدد من المفردات.

أسفرت النتائج أن الاختبار التواؤمى المحوسب أكثر فعالية من الاختبار الخطى المحوسب. حيث تمكن من تحقيق دقة قياس measurement Precision أعلى من خلال تطبيق (٢٥) مفردة فى المتوسط بينما كان عدد مفردات الاختبار الخطى (٦٠) مفردة . بل إنه بلغ دقة أعلى مع تطبيق (١٣) مفردة فقط، ووصل إلى معامل ثبات تجاوز (٠,٨٠) من خلال تطبيق (٥) مفردات فقط كما أمكن للاختبار التواؤمى المحوسب الوصول إلى معامل ثبات (٠,٩٤) من خلال تطبيق (٥٠٪) من عدد المفردات فى الاختبار الخطى.

كما أسفرت النتائج ايضا أن أسلوب تقدير القدرة البعدى الأقصى أكثر دقة من أسلوب الأرجحية القصوى لتقدير القدرة حيث أنه يقدم معلومات Information أكثر وله من الخطأ المعيارى Standard Error أقل بدلالة. غير أن الارجحية القصوى أعطى متوسطا لتقدير القدرة أعلى بدلالة من البعدى الاقصى. وأثبتت الدراسة ان الاختبار التواؤمى المحوسب يمتاز فى أنه: يوفر فى عدد مفردات الاختبار بنسبة تزيد عن (٥٠٪) فى المتوسط عن الاختبار الخطى. ويزيد من دقة القياس بنسبة يصل متوسطها إلى (٢٠٪) عن الاختبار الخطى.

ويعطى مقدارا من المعلومات أعلى بنسبة تتجاوز (٣٢٪) فى المتوسط عن الاختبار الخطى. ويعطى تقديرا للقدرة لا يختلف متوسطه عن متوسط تقدير القدرة للاختبار الخطى.

(٣) **اختبار رسم الرجل:** أعدت هذا الإختبار الباحثة الأمريكية جودانف Goodenough عام ١٩٢٦ ، ويتطلب من المفحوص باختصار «رسم صورة لرجل»، وقد ظل هذا الاختبار فى ميدان الاستخدام العام دون تعديل منذ ظهوره حتى عام ١٩٦٣ حيث ظهر تعديل جديد شامل له باسم «اختبارالرسم لجودانف وهاريس Goodenough-Harris Drawing Test ويؤكد هذا التعديل كما يؤكد الاختبار الأصلى - دقة الطفل فى الملاحظة وارتقاء تفكيره المجرد، وليس المهارة الفنية فى الرسم، حيث تعطى درجة لكل جزء من الجسم رسمه الطفل، وتفاصيل الملبس، والنسب، والمنظور، وغير ذلك من الخصائص التى بلغت فى طبعة ١٩٦٣ (٧٣ وحدة قابلة للقياس)، وقد اختيرت هذه الوحدات على أساس تمايز العمر، والعلاقة بالدرجة الكلية فى الاختبار، والارتباط بدرجات اختبارات الذكاء الجماعية الأخرى. وتم الحصول على ذلك من اختبار ٥٠ طفلا ، و٥٠ طفلة فى كل مستوى دراسى ابتداء من الحضانة حتى سن ١٥ سنة. وقد أجريت دراسات كثيرة على المقياس المعدل ثبت منها أنه ثابت وصادق، وأنه إختبار غير لفظى للنضج العقلى بالإضافة إلى حساسيته الشديدة للعوامل الثقافية التى تؤثر فى تمثيل الطفل لمفهوم شكل الإنسان، ومعنى ذلك أن هذا الاختبار ليس متحررا من أثر للثقافة كما كان مفترضا.

ويتطلب المقياس المعدل أن يرسم الطفل أيضًا صورة لامرأة، ويصحح هذا المقياس الجديد الخاص برسم المرأة فى ضوء ٧١ وحدة تتشابه مع مقياس رسم الرجل، كما يطلب من الطفل أن يرسم نفسه Self Scale ولو أن هذا المقياس الثالث أعد فى جوهره كأسلوب إسقاطى لدراسة الشخصية، والبيانات المتوافرة عنه ليست مقنعة، وأعدت معايير لمقياس رسم الرجل ومقياس رسم المرأة من تطبيقه فى أمريكا على ٣٠٠ طفل من كل مستوى عمرى ابتداء من سن ٥ سنوات حتى سن ١٥ سنة. ويمكن تحويل الدرجات الخام فى كل مقياس إلى درجات معيارية معدلة متوسطها ١٠٠ وانحرافها ١٥ ، وحسب له الثبات بطريقة إعادة الاختبار وطريقة التجزئة النصفية، واستخدم فى حساب صدقه محك تمايز العمر ومحك التنبؤ بالنجاح المدرسى، كما استخدم فى بعض الدراسات مقياس ستانفورد - بينيه كمحك أيضًا، وكذلك تقديرات المعلمين وتتابعت الدراسات السيكومترية حول الاختبار فى مختلف الثقافات فأكدت الدراسات الأمريكية اللاحقة جميعا ثبات الاختبار وصدقه، وقد قنن

الاختبار فى البيئة المصرية وعدد من البيئات العربية الأخرى.

وقد تكون أولى الدراسات التى أجريت فى هذا الصدد دراسة بإشراف القوصى (٥٥) والتى قام بها ميشيل اسكندر (٥٥) بين عامى ١٩٣٤، ١٩٣٧ بتطبيق الاختبار على عينة مصرية اقتربت من الألفين، أعمارهم بين ٤، ١٣ سنة، ووضع نموذجا وسطا لكل سن كما أعطى معيارا مصريا للعمر العقلى، وتحدد صدق القياس باستخدام محك اختبار الذكاء الابتدائى للقبانى، كما تمت المقارنة بين الصفوف الدراسية وبين الأعمار وبين الأطفال العاديين والمتخلفين وفى جميع الحالات كان التمييز فى الاتجاه المتوقع على نحو يدل على صدق التكوين الفرضى، ولم يشر البحث إلى حساب الثبات، وبعد ذلك قام على شلتوت بتعديل طريقة تصحيح الاختبار بما يلائم البيئة المصرية، وظل الاختبار يستعمل فى العيادة النفسية بكلية التربية جامعة عين شمس منذ ذلك الحين.

وقد قام مصطفى فهمى (٥٣) بدراسة على هذا الاختبار حيث طبق على ٩٧٦ تلميذا وتلميذه فى فئات عمرية من ٦ سنوات حتى ١٢ سنة حسب من درجاتها معايير الأعمار العقلية مستخدمًا وحدة عمرية طولها سنة كاملة، ويلاحظ على هذا البحث بعض المشكلات المنهجية منها استخدام معادلة غير واضحة لحساب ثباته بطريقة كيودر- ريتشاردسون، فقد أشار فى مراجعة إلى معادلة معينة واستخدم فيما يبدو معادلة أخرى، بالإضافة إلى الاعتماد على ٤٠ مفردة فقط، بينما العدد الأصلى للمفردات - كما حددته جودانف فى الاختبار الأصلى الذى اعتمد عليه فى هذا البحث - هو ٥١ مفردة. كما لوحظ عدم وضوح محك التمايز العمر المستخدم فى تقدير الصدق.

ومن الدراسات المهمة التى أجريت على هذا الاختبار تلك التى قام بها الباحث السودانى مالك البدرى (٤٤أ فى ٤٥) وفيها لخص البحوث العربية السابقة التى أجريت حول الاختبار ومنها رسالة الماجستير التى تقدم بها محمد بطانية إلى الجامعة الأمريكية فى بيروت، حيث طبق الإختبار على ٥٠٠ طفل أردنى تراوحت أعمارهم بين ٥ و ١٠ سنوات، وحسب الثبات بطريقة إعادة الاختبار (ن = ٣٠) فبلغ ٠,٨١ واستخدم محك تمايز العمر فى تقدير الصدق فوجد أن الفروق بين متوسطات الأطفال

التقويم النفسى ـــــــ

العناصر التى اقترح الباحث إضافتها رسم الشارب أواللحية أو كليهما، إظهار الوشم، تظليل الجسم، رسم العصا فى اليد. ويبدو أن نتائج مالك البدرى تصلح لكثير من الدول العربية التى يرتدى الرجال فيها «الجلابية»، أو «الدشداشة» أو «الثوب» كالسعودية والكويت ومنطقة الخليج، وبعض أنحاء مصر والمغرب العربى.

وفى الكويت أجرى محمد نسيم رأفت (٥٢) دراسة على الاختبار تهدف إلى تحديد نوعين من المعايير أحدهما يناسب الزى الأوربى وثانيهما الزى التقليدى الكويتى وقد أجريت الدراسة على ٢٥٦٢ طفلا وطفلة تمتد أعمارهم بين ٤، ١٢ سنة.

وفى لبنان قام عبدالسلام عبد الغفار والسيد فؤاد أعظمى بتطبيق الاختبار على عينة مكونة من ١٨٥٥ طفلا وطفلة تتراوح أعمارهم بين ٦، ١٢ سنة واستخدم فى تصحيح الاختبار ٥٠ مفردة فقط من مفتاح جودانف الأصلى، وحسب الثبات بطريقة إعادة الاختبار فبلغ الارتباط ٠,٧٥ وفى حساب الصدق استخدم محك تمايز العمر، وأعدت معايير أعمار عقلية.

وفى اليمن قام فتحى السيد عبد الرحيم (٣٦) بتطبيق الاختبار على أطفال منطقة صنعاء من ذوى الأعمار الزمنية التى تمتد من ٦ – ١٢ سنة بلغ عددهم ١٤٨٢ طفلا وطفلة، وحذا الباحث فى بحثه حذو بحث مصطفى فهمى السابق الإشارة إليه، فاستخدم مفرداته الأربعين وحسب معايير العمر لوحدة زمنية طولها عام كامل واستخدم طريقة كيودر- ريتشاردسون فى حساب الثبات، وفى حساب الصدق استخدم محك تمايز العمر فكانت الفروق دالة بين الأعمار، كما حسب معامل الارتباط بين الاختبار ومحك تقديرات المعلمين فبلغ ٠,٥٩ وهومعامل مرتفع نسبيا بالمقارنة بالمعاملات التى لوحظت فى البحوث السابق الإشارة إليها.

وقد أجريت هذه الدراسات على اختبار جودانف الأصلى طبعة ١٩٢٦ أما بالنسبة للاختبار المعدل لجودانف وهاريس الصادر عام ١٩٦٢ فلم يجر عليه إلا عدد محدود من الدراسات فى الثقافتين المصرية والعربية، فقد قام محمد فرغلى فراج وعبد الحليم محمود السيد وصفية مجدى بترجمة كراسة تعليمات هذا الاختبار عام ١٩٧٦، وفى العام نفسه أجرى محمد متولى غنيمه (٥٢) تقنين هذا المقياس الجديد على البيئة المصرية حيث طبق الاختبار على ١٩١٥ طفلا وطفلة من سبع مدارس ابتدائية

بمحافظة القاهرة، وعلى الرغم من أن بعض الأطفال رسموا الرجل مرتديا الجلباب إلا أن الباحث اعتمد فى التحليل على أولئك الذين رسموا الرجل مرتديا الزى الأوربى مستخدما مفتاح هاريس المعدل الذى يتألف من ٧٣ مفردة،، وحسب الثبات بطريقة تكرار المصحح الواحد، وتكرار المصححين المختلفين، فبلغ معامل الموضوعية فى الحالة الأولى ٠,٩٣ ، وفى الحالة الثانية بين ٠,٦١ ، ٠,٨٧ ، وكذلك استخدم طريقة إعادة الاختبار فبلغ معامل الثبات ٠,٨٤ ، ٠,٩٨ للأعمار المختلفة، وفى حساب الصدق المرتبط بالمحكات استخدم الباحث مقياس ستانفورد - بينيه وبلغ معامل الارتباط بين ٠,٨٠ ، ٠,٥٤ ، والتحصيل الدراسى وكانت معاملات الارتباط بين ٠,٤٢ ، ٠,٨١ ، كما قارن الباحث بين مجموعتين متضادتين أحداهما من المتخلفين عقليا والثانية من الأطفال العاديين وكانت الفروق دالة فى الاتجاه المتوقع، كما استخدم الباحث محك تمايز العمر ، ولم يظهر البحث فروقا دالة بين الجنسين فى معظم الصفوف الدراسية.

ولعل من المشكلات الجوهرية التى يظهرها هذا الاختبار مسألة زى الرجل المرسوم، وهى مشكلة حاول أن يحلها كل من مالك البدرى ومحمد متولى غنيمة عن طريق التضمين أى التساهل فى تصحيح بعض المفردات. وقد قامت آمال أحمد مختار صادق (١٠) فى عام ١٩٧٩ بدراسة مقارنة مستقلة للمقارنة بين رسوم الرجل مرتديا الزى التقليدى السعودى والزى الأوربى، وقد أجريت الدراسة على عينة من الأطفال قوامها ١٩٣٣ طفلا وطفلة تمتد أعمارهم بين ٨ سنوات، ١٥ سنة، وقد اختارت سن ٨ سنوات كحد أدنى لأنه العمر الذى أظهرت البيانات الامبريقية أن نوعى الزى يتمايزان فيه فى رسوم الأطفال، وقامت الباحثة بدراسة الفروق بين النسب لكل مفردة من مفردات المقياس فى نوعى الرسم وكانت الفروق الدالة جميعا لصالح الزى الأوربى، وتعلق الباحثة على هذه النتيجة بقولها (١٠: ١١٩):

«تؤكد هذه النتيجة أن رسم الرجل الأوربى يهيىء الفرصة للطفل أن يتنبه إلى تفاصيل أكثر لا ترجع إلى طبيعة الزى وحده.. فكثير من المفردات التى أظهرت فروقا دالة لصالح الزى الأوربى تتضمن تفاصيل الجسم البشرى ذاته ومن ذلك وجود الرقبة ووجود الشعر».

ومن الدراسات التقنية المهمة التى أجريت على اختبار جودانف ـ هاريس طبعة ١٩٦٣ تلك التى أجريت فى المملكة العربية السعودية عام ١٩٧٧ وقام بها فريق من الباحثين باشراف فؤاد أبو حطب (٢٧) ولعل أهم الإضافات فى هذا البحث أن بروتوكولات التصحيح المستخدمة (وهى من إعداد فؤاد أبو حطب) جعلت عدد المفردات ٧٧ مفردة نتيجة إضافة مفردات جديدة أومفردات تضمنها الاختبار الأصلى (طبعة ١٩٢٦)، وقد أجرى البحث على ٢١٥ طفلا وطفلة من مرحلة ما قبل المدرسة تمتد أعمارهم بين ٣ – ٥ سنوات، ١١٩١ طفلا وطفلة تمتد أعمارهم بين ٦ – ١٥ سنة اختيروا جميعا من رياض الأطفال والمدارس الابتدائية والمتوسطة بالمنطقة الغربية بالمملكة العربية السعودية، وحسبت معاملات شيوع المفردات التى أظهرت أن عدد المفردات التى تعد أفضل مقياس للذكاء باستخدام رسم الرجل مرتديا ثوبا (الزى الوطنى السعودى) هو ٦٦ مفردة، أما باقى المفردات فهى أكثر ملاءمة لرسم الرجل مرتديا الزى الأوربى.

وفى حساب صدق المفردات استخدمت طريقة المقارنة الطرفية وحسبت معاملات الارتباط الثنائية بين المفردة والدرجة الكلية فى الاختبار، أما بالنسبة لصدق الاختبار ككل استخدم محك تمايز العمر فى تقدير صدق التكوين الفرضى، والأداء فى اختبار المصفوفات المتتابعة (والذى سبق تقنينه على نفس الثقافة) لتقدير الصدق المرتبط بالمحكات، أما عن ثبات الاختبار فقد حسب بطريقة إعادة الاختبار وبطريقة كيودر ـ ريتشاردسون ، وكانت المعاملات ـ كما هو متوقع ـ منخفضة فى الأعمار الدنيا ـ وخاصة بالنسبة لطريقة إعادة الاختبار أقل منها بالنسبة لطريقة كيودر ـ ريتشارد سون، وهى نتيجة متوقعة فى ضوء طبيعة كل من الطريقتين فى حساب الثبات، وتباين الخطأ المتضمن فى كل منهما.

وتحتل مسألة الموضوعية أهمية خاصة فى اختبار رسم الرجل، وقد حسبت بالطرق الثلاث الآتية:

١ – إعادة المصحح الواحد باستخدام بروتوكول التصحيح فى إحدى المرتين وبدون استخدامها فى المرة الثانية.

٢ – إعادة المصحين باستخدام بروتوكولات التصحيح بالنسبة لبعض المصححين وبدون استخدامها بالنسبة للبعض الآخر.

٣ – إعادة المصحح الواحد باستخدام بروتوكولات التصحيح فى الحالتين.

وكانت معاملات الموضوعية مرتفعة، إلا أنها كانت منخفضة فى حالة الاختلاف فى استخدام البروتوكولات أو عدم استخدامها.

وحسبت للاختبار معاييره فى صورة نسب ذكاء انحرافية متوسطها ١٠٠ وانحرافها المعيارى ١٥ لجميع الأعمار.

وفى عام ١٩٨٢ قام نعيم عطية (٥٨) بمحاولة بناء نسق نظرى للاختبار له طبيعة هرمية يتألف من عدة عمليات عقلية يحتاجها الأداء فى الاختبار وتشمل: ثبات الصورة الحسية، تركيز مواقعها، إدراك الحجم النسبى، إدراك التناسب الداخلى، إدراك أبعاد التجسيم، الاتساق البصرى ـ الحركى، توجيه الخطوط ، الإيماء والتعبير، الاستكمال، الرسم الجانبى، ويقترح ثلاث مراحل لإجراء الاختبار وهى:

١ – **الرسم المواجه** ، وفيها يطلب من الطفل رسم صورتين للرجل فى وضع مواجه أحدهما تلقائية والثانية فى ضوء تنبيهات معينة.

٢ – **استقصاء الرسوم** واستكمالها وتتألف من خطوتين: الاستقصاء لاستدراج التسميات والاستقصاء لاستكمال النواقص.

٣ – **الرسم الجانبى.**

ويصف الباحث بالتفصيل بروتوكولات التصحيح المقترحة لعدد من المفردات بلغ ٢٢٦ مفردة، ولم يجر الباحث تقنينا شاملا للمقياس بسبب ظروف لبنان ـ وطنه، واكتفى بدراسة إمكانات التعديل الذى يقترحه للتقنين بأن طبقه على عينة صغيرة نسبيا (ن = ٣٢٩) من أطفال تمتد أعمارهم بين ٤,٥ سنة، ٨ سنوات، وحسب الثبات بإعادة الاختبار وبلغ معامل الارتباط ٠,٧١ (ن = ٥٥)، واستخدم فى حساب الصدق محك تمايز العمر (أو ما يسميه الباحث تدرج النتائج مع السن)، أما المعايير فاتخذت صورة نسب الذكاء التقليدية بمتوسط ١٠٠ وانحراف معيارى ١٦.

٣ – اختبار هايدلبرج للذكاء:

هذا الاختبار أعده فى أصله الألمانى كراتزيمر Kratzmeier عام ١٩٧٧ ونقله إلى العربية حسنين الكامل (٢٣) عام ١٩٨٥، ويقوم فى جوهره على تحليل ما يسمى

«القدرة التمييزية» إلى خمسة مكونات هى: إدراك العلاقات، والسلاسل المنطقية، والإنتاجية والتمييز والابتكارية، وقد استبعد معد الاختبار بالعربية القسم الخامس (الابتكارية) من الصورة العربية، وطبقه على عينة مكونة من ٢١٦ طفلا وطفلة من تلاميذ الصفوف الثانى والثالث والرابع من مرحلة التعليم الأساسى بمحافظة سوهاج، وأجرى على كل اختبار فرعى (يتألف من ٧٢ مفردة) تحليلا للمفردات بحساب معاملات الصعوبة والتمييز ثم طبق الاختبار على عينة أخرى (ن = ٨٢٤) لها نفس الخصائص السابقة وحسبت لها المعايير فى صور مئينيات ثم حسبت نسب الذكاء الانحرافية، بمتوسط ١٠٠ وانحراف معيارى ١٥ والدرجات التائية.

وحسب الثبات بمعامل ألفا لكرونباك، أما صدق التكوين الفرضى فقد حسب بطريقتين،: أولاهما تمايز العمر، وكانت الفروق فى الاتجاه المتوقع، والطريقة الثانية هى التحليل العاملى وتوصل الباحث إلى أربعة عوامل تطابقت مع الاختبارات الفرعية فيه، وحسب الصدق التلازمى باستخدام محك التحصيل المدرسى وكانت معظم معاملات الارتباط دالة، وباستخدام اختبار الذكاء المصور كمحك وكانت معاملات الارتباط منخفضة بشكل واضح يوحى بضرورة إعادة تقنين اختبار الذكاء المصور ذاته وإعادة النظر فيه كمقياس للذكاء.

نحو مفهوم شامل للذكاء (*)

تضمينات النموذج الرباعى للعمليات المعرفية

مع ظهور النماذج المعرفية وتطورها خلال السنوات الأخيرة ظهرت الحاجة إلى تصنيف نتائج البحوث إلى فئات من العمليات المعرفية. وكان النموذج الذى اقترحه فؤاد أبو حطب فى كتاب القدرات العقلية منذ طبعته الأولى عام ١٩٧٣ والطبعات المتتالية (١٩٧٨ ، ١٩٨٠، ١٩٨٣، ١٩٩٦) احدى المحاولات فى هذا الاتجاه حيث استمر يطور نموذجه طوال عقد الثمانينات وخلال عقد التسعينات من هذا القرن العشرين حتى وصل إلى الصورة التى نعرضها فى الدراسة الحالية.

وقد آثر فؤاد أبو خطب عند عرض نموذجه، عرض نموذج هوارد جاردنر الذى

* عن الدراسة التى قدمها أ.د. فؤاد أبو حطب فى مؤتمر العمل المصغر الذى عقده المركز القومى للإمتحانات والتقويم التربوى لإعداد مقياس قومى للذكاء (نوفمبر ١٩٩٥) وهى دراسة مؤسسة على الفصول (٨،١٧، ١٨، ١٩، ٢٠، ٢١) من كتاب القدرات العقلية ١٩٩٦ .

نشر لأول مرة عام ١٩٨٣، ثم عرض لنموذجه الرباعى العملياتى فى مراحل تطوره المختلفة ليبين للقارئ أوجه الشبه والاختلاف بين النموذجين، وتوثيقاً لحقه التاريخى فى السبق التصنيفى حتى لا تختلط الأوراق أمام القارئ العربى فى جو ثقافى عام لا يدرك الإبداع إلا فيما ينتجه الغرب، ولايدرك فى الإنتاج العلمى العربى إلا التبعية والمحاكاة.

نموذج الذكاء المتعدد عند جاردنر

ظهر النموذج لأول مرة فى كتاب صدر للعالم الأمريكى المعاصر هوارد جاردنر عام ١٩٨٣ (Gardner, 1983) (*) عنوانه «أطر العقل». ومنذ ذلك الوقت حظى الكتاب ومؤلفه بشهرة فائقة تجاوزت حدود ما أنجزه بالفعل. ونعرض فيما يلى المعالم الأساسية لهذا النموذج.

المعالم العامة للنموذج:

يختلف نموذج جاردنر عن النماذج السابقة فى أنه لا يستند إلى محك واحد مثل مفهوم معامل الارتباط بين الاختبارات كما هو الحال فى نماذج التحليل العاملى، أو أساس نتائج البحث التجريبى المعملى فى علم النفس كما هو الحال فى نماذج تجهيز المعلومات، وإنما يقترح وجود أنواع متعددة مستقلة من الذكاء. وكل نوع منها لابد أن تتوافر فيه ثمانية محكات تؤيد وجوده يحددها جاردنر على النحو الآتى:

(١) إمكانية تحديد نوع الذكاء فى ضوء نتائج الدراسات فى اصابات المخ. وهذا الدليل يعتمد على نتائج البحث فى ميدان علم النفس العصبى التى تؤكد وجود بنى عصبية مختلفة لكل نمط من الذكاء فى ضوء تأثير المواضع المصابة فى المخ على أنشطة عقلية معينة.

(٢) ظاهرة الأفراد غير العاديين من نوع الحكماء المعتوهين والعباقرة وغيرهم، والتى تؤكد النمو المستقل لأنماط مستقلة من الذكاء الإنسانى.

(٣) وجود مجموعات أو محاور معينة من العمليات ترتبط بالميكانيزمات العصبية، ويختلف عدد هذه المجموعات أو المحاور من نمط للذكاء إلى آخر. فالذكاء

* المصادر الواردة فى هذه الدراسة يمكن الرجوع إليها فى الكتاب الأصلى للمؤلف : Gardener, H., Frames of the mind : The theory of multiple intelligence, London- Heine mann, 1983.

الموسيقى مثلاً يعتمد على محور واحد للعمليات هو القدرة على تمييز الدرجة الصوتية. أما الذكاء اللغوى فيعتمد على أربعة محاور هى: (أ) القدرة البلاغية أى القدرة على اقناع الآخرين، (ب) القدرة التذكرية، (ج) القدرة الشارحة، (د) القدرة على توضيح المعانى.

(٤) وجود تاريخ نمائى مميز مستقل لكل نمط من أنماط الذكاء فى مراحل العمر المختلفة مع وجود مجموعة محددة من الأداءات تميز مستوى الخبرة العالية فيه.

(٥) وجود تاريخ تطورى لكل نمط من أنماط الذكاء، يرتبط بالصور البدائية والمكبرة للتعبير عنه لدى الأنواع الحيوانية الأخرى.

(٦) توافر أدلة من نتائج البحث فى علم النفس التجريبى. ومن أمثلة ذلك ما أكدته بحوث علم النفس المعرفى حول المهام المزدوجة أنه حين تتطلب هذه المهام نفس النمط من الذكاء يزداد بينها التداخل اذا قورنت بالمهام التى تتطلب انماطا مختلفة من الذكاء.

(٧) وجود تأييد لنمط الذكاء من أدلة البحث ذات المنحى السيكومترى الذى يعتمد على أساليب التحليل العاملى حيث فئات الاختبارات التى تقيس أنواعاً مختلفة من الذكاء تكون معاملات الارتباط بينها أقل من تلك التى تقيس نفس النمط من الذكاء.

(٨) قابلية نمط الذكاء للتشفير فى نسق رمزى معين تحدده الثقافة التى يعيش فيها الأفراد فالذكاء اللغوى يتم تشفيره فى صورة لغوية لها قواعدها، والذكاء الموسيقى يتم التعبير عنه فى لغة صوتية.

- أنماط الذكاء:

يقترح جاردنر فى كتابه الأصلى الصادر عام ١٩٨٣ ستة أنواع مستقلة من الذكاء هى الذكاء اللغوى والذكاء الموسيقى والذكاء المنطقى الرياضى والذكاء المكانى والذكاء الجسمى الحركى والذكاء الشخصى. وكانت صيغته المبكرة للذكاء الشخصى تشمل كلاً من الذكاء الشخصى والذكاء الاجتماعى. إلا أنه فى عام ١٩٨٦ فصل بين النوعين فى مقال مشترك مع وولترز (Walters & Gardener, 1986)وأصبح

عدد أنماط الذكاء المستقلة سبعه يسميها الذكاءات المتعددة Multiple Intelligences (MI) هى كما يلى:

١- الذكاء الموسيقى: فالقدرة الموسيقية التى عوملت على أنها قدرة خاصة ولم تظهر بصورة واضحة فى مقاييس الذكاء العام، يرى جاردنر أنها تؤلف نمطا مستقلا للذكاء. ويذكر أدلة نمائية ونيرولوجية وثقافية على ذلك.

ومن الأدلة النيرولوجية ارتباط الذكاء الموسيقى بالنصف الكروى الأيمن بالمخ، على الرغم من أنه لا يوجد له موضع واضح ومحدد فى المخ كما هو الحال فى اللغة مثلاً.

٢- الذكاء الجسمى - الحركى: ويشير إلى القدرات التى تتطلب الاستخدام البدنى الصريح سواء للجسم ككل أو لبعض أجزاء منه فى حل المشكلات أو إنتاج النواتج أو العروض فى أنشطة مثل الرقص والألعاب الرياضية والتمثيل والعزف على الآلات الموسيقية والجراحة فى الطب. ومرة أخرى يحدد الأساس النيرولوجى لهذا الذكاء فى اللحاء الحركى مع سيطرة النصف الكروى للمخ على نصف الجسم الذى يرتبط به. وعلى الرغم من أهمية هذا النوع من الذكاء أهملته اختبارات الذكاء الحالية.

٣- الذكاء المنطقى والرياضى: وهذا الذكاء هو ما شاعت تسميته بالتفكير العلمى. وفى رأى جاردنر أنه مستقل عن الذكاء اللغوى لأن حل المشكلة قد يتوصل إليه الباحث العلمى قبل صياغته لفظيا. وتلعب فيه القدرات الاستدلالية بالإضافة إلى مهارات الملاحظة دوراً واضحاً. وهذا النوع من الذكاء له موضع أساسى فى معظم اختبارات الذكاء الراهنة.

٤- الذكاء اللغوى: وهو اكثر أنواع الذكاء شيوعاً فى الاختبارات التقليدية. وتوجد مواضع فى المخ مسئولة عنه فى النصف الكروى الأيسر من المخ وعلى وجه الخصوص «منطقة بروكا» المسئولة عن إنتاج البنية النحوية للغة. ويتجاوز هذا النوع من الذكاء حدود المدخلات السمعية والمخرجات الرمزية المعتادة، كما هو الحال عند الأطفال الصغار، والصم من الكبار حيث تستخدم لغة الاشارة.

٥- الذكاء المكانى: يرى جاردنر أن القدرة المكانية تستحق فى ضوء المحكات التى يستخدمها أن تؤلف فئة مستقلة من الذكاء. صحيح أن حظ هذه القدرة كان

أفضل كثيراً من القدرة الموسيقية فى احتواء اختبارات الذكاء الحالية على أمثلة لها، إلا أنها ذات أهمية خاصة لدى الملاحين فى البحر والجو، ولدى الجغرافيين وغيرهم من مستخدمى الخرائط. ويلعب التصور البصرى المكانى – وهو أحد مكوناتها – دوراً هاماً فى الفنون البصرية.

وتؤكد الأدلة النيرولوجية أن هذا النوع من الذكاء يعتمد على نشاط النصف الكروى الأيمن. ولا يعتمد الذكاء المكانى على الادراك البصرى فقط، فعند المكفوفين يساعدهم الادراك اللمسى على تحديد المواضع والمسافات والحجوم. ويرى جاردنر أن الاستدلال المكانى عند المكفوفين يحل محل الاستدلال اللغوى عند المبصرين.

٦– الذكاء الاجتماعى: ويسميه جاردنر الذكاء بين الأشخاص interpresonal والذى يعتمد على ادراك الفروق بين الآخرين وخاصة ما يتصل بدوافعهم ومقاصدهم وحالاتهم الوجدانية والمزاجية. ويتمثل هذا النوع من الذكاء فى قدرة الشخص على «قراءة» رغبات ومقاصد الآخرين حتى ولو لم تكن واضحة. ويظهر هذا الذكاء فى سلوك رجال الدين والقادة السياسيين والمعلمين والمعالجين والآباء والأمهات. ولا يعتمد هذا النوع من الذكاء على اللغة.

٧– الذكاء الشخصى: ويسميه جاردنر الذكاء داخل الفرد intrapersonal والذى يعتمد على معرفة الجوانب الداخلية فى الفرد نفسه، وإدراك الفرد لمشاعره وانفعالاته، وقدرته على التمييز بينها وتصنيفها واستخدامها فى فهمه لسلوكه هو وتوجيه هذا السلوك.

ويوضح جاردنر أن كلاً من الذكاء بين الأشخاص و الذكاء داخل الشخص الواحد – واللذين كان يسميهما الذكاء الشخصى فى كتابه المبكر لم يحظ أى منهما باهتمام يذكر فى اختبارات الذكاء التقليدية.

تقويم نموذج جاردنر:

يذكر برودى (Brody,1992)أن من الصعب تقويم نموذج جاردنر لأنه لا يكاد يذكر أى دراسات منظمة أجريت لاختبار فروض مشتقة منه، ولعل أهم هذه الفروض هو استقلال أنماط الذكاء التى يتضمنها هذا النموذج، ناهيك عن مدى توافر جميع المحكات الثمانية التى أشار إليها فى كل نمط من هذه الأنماط.

التقويم النفسى ـــــــ

يذكر جاردنر فى كتابه (Gardner,1983)أن هذه المحكات اعتباطية وتعتمد فى جوهرها على التقدير الاجتماعى لأنواع معينة من الذكاء الإنسانى. فإذا كان الأمر كذلك فلابد من توافر أدلة علمية على أن هذه الأنماط التى اختارها جاردنر تحظى بالتقدير الاجتماعى بالفعل، فالاعتماد فى مثل هذا الحالة على الانطباع الشخصى وحدة لا يكفى.

ويبقى السؤال الأخير: هل المحكات التى اعتمد عليها جاردنر فى تحديد أنماط الذكاء تدعم قائمته المقترحة؟ للإجابة على هذا السؤال نركز على المحكات النيرولوجية. وفى هذا نذكر الحقيقتين الآتيتين:

(١)- تؤكد بحوث إصابة المخ وجود أنواع أخرى من الذكاء غير تلك التى اقترحها جاردنر. ومن أمثلة ذلك الاصابة الثنائية الطرف الذى تتعرض لها اللحاءات البصرية ويترتب عليها الخلل الإدراكى المسمى Prosopognosia. وهذا الخلل يتسم بعدم القدرة على التعرف على الوجوه المألوفة، والذى قد يتخذ صورة عامة فى عدم القدرة على التعرف على الوحدات التى تتألف منها فئة من الأشياء التى تدرك بصرياً. وهى قدرة لها أهميتها الاجتماعية فى بعض الثقافات كالثقافة الأفريقية والبدوية حيث يكون على الفرد معرفة كل حيوان فردى باستخدام منبهات معينة (فى قطيع الغنم أو الماشية الذى يرعاه)، وهو محك ثقافى من بين المحكات التى يعتمد عليه جاردنر فى اختياط أنماط الذكاء.

ولعلنا نذكر القارئ بأن هذه القدرة ترتبط بعوامل الوحدات عند جيلفورد. ويمكن أن تتوافر أدلة نيرولوجية وثقافية على أنماط أخرى من الذكاء لم تشملها قائمة جاردنر.

(٢)- لا تقدم أدلة النيرولوجية أدلة واضحة على استقلال أنماط الذكاء. وقد اعترف جاردنر نفسه بهذه الحقيقة (Gardner,1983)حين أشار إلى أن الذكاء المنطقى الرياضى لا يتأثر بخلل موضعى فى المخ وإنما من اضطرابات عامة مثل الخرف Dementiaحيث تختل مساحات كبيرة من الجهاز العصبى بسرعة واضحة. ويذكر (Brody,1992)أن هذه الاشارة من جاردنر تصل نموذجه بالنماذج الهرمية التى اقترحها بعض أصحاب التحليل العاملى. وبالطبع فان القدرات

الأقل اعتماداً على التموضع المكانى فى المخ هى الأقرب إلى العامل العام، وهو التفسير الذى قدمه ثرستون مبكراً عام ١٩٤٨ للعامل العام من الدرجة الثانية (Thur-stone,1948).

ولعل هذا كله ما دفع فؤاد أبوحطب إلى تعديل نموذجه الذى اقترحه لأول مرة عام ١٩٧٣ عند صدور الطبعة الأولى من كتاب القدرات العقلية وقد بدأ هذا النموذج باقتراح سبعة أنواع من الذكاء تتضمن معظم الفئات التى تناولها جاردنر فى كتابه بعد عشر سنوات (Gardner,1983)ثم تطور هذا النموذج طوال السنوات العشرين الماضية إلى صورته الحالية التى نعرضها فى القسم التالى.

النموذج الرباعى العملياتى عند فؤاد أبو حطب

– لماذا الحاجة إلى نموذج جديد للقدرات العقلية؟

بدأ اهتمام المؤلف بالعمليات المعرفية (أو ما سمى فيما بعد بعلم النفس المعرفى) منذ ايام دراسته العليا بجامعة لندن للحصول على الماجستير والدكتوراه فى علم النفس خلال الفترة ١٩٦١ – ١٩٦٦، ثم ظهر صريحاً عقب عودته من البعثة فى مطلع عام ١٩٦٧ حين طلب منه المرحوم أ.د. أحمد زكى صالح رئيس القسم الذى عاد إليه أن يقترح مقرراً لطلاب الدراسات العليا فأجاب بأن يكون مقرراً فى العمليات المعرفية. وقد كان وظل يدرس هذا المقرر دون انقطاع إلا لضرورة – منذ ذلك الوقت. ومنذ ذلك الحين تنبهنا إلى حقائق ثلاث:–

١– فشل النماذج التصنيفية للقدرات العقلية التى كانت قائمة فى ذلك الوقت فى تحديد ما كانت تشير إليه باسم العمليات العقلية. وكل التصنيفات المقترحة ابتداء من تشارلز سبيرمان وحتى جو جيلفورد وبنجامين بلوم كانت من نوع التنظير التاملى.

٢– معظم النقص يعود فى جوهره إلى تبنى تعريف إجرائى «للقدرة» يعتمد فى جوهره على الأساليب التقليدية للتحليل العاملى، وهو التعريف الشهير الذى صاغه فيليب فرنون منذ عام ١٩٥٠ فى كتابه الشهير «بنية القدرات الانسانية» ومؤداه أن القدرة هى مجموعة من أساليب الأداء (أو الاختبارات) ترتبط فيما بينها ارتباطاً موجباً عالياً، وترتبط بغيرها من أساليب الأداء ارتباطاً منخفضاً. وهذا كما يبدو للمهتم بالتحليل العاملى هو التعريف المناسب للعامل ذى البنية البسيطة كما يحدده ثرستون،

إلا أن هذا التعريف يتضمن أن القدرة، كتكوين فرضى، إنما تستنتج من المتغيرات التابعة أى أساليب الأداء وحدها. أو هى تنتمى إلى نمط السيكولوجية الذى يسميه س.س. ستيفنس «سيكولوجية س- س».

٣- هذا المنحنى الجامد من التحليل العاملى أدى إلى خلق نوع من الصراع مع المنحى التجريبى التقليدى، والذى يصنفه ستيفنس فى فئة سيكولوجية م- س. وكان أول من تنبه إلى خطر هذا الصراع كرونباك الذى نشر فى عام ١٩٥٧ مقاله الشهير الذى دعا فيه إلى إحداث لون من التآزر بين علم النفس الفارق وعلم النفس التجريبى. ومع جدية دعواه إلا أن الصراع لم يتوقف أبداً، على الرغم من تغير مظاهره، والتى لا تزال تتغير، كما هو الحال اليوم بين الاتجاه السيكومترى واتجاه تجهيز المعلومات.

ولعل أشد صور الخلاف حدة ما اتصل فى جوهره بتصنيف العمليات النفسية على وجه العموم. فأسس التصنيف فى معظم الحالات غامضة وتعتمد فى أغلب الأحوال على أسس تأملية محضة. ولعلنا نشير هنا إلى أن جونسون (Johnson 1954)يذكر قائمة تتألف من ٥٤ مفهوماً باللغة الإنجليزية أشير إليها فى مختلف البحوث والكتابات المتصلة بالنشاط المعرفى للإنسان، بل والمستخدمة أيضاً فى لغة الحياة اليومية. ومعظم البحوث التى تناولت العمليات المعرفية بالتحليل عالجت بعض هذه المفاهيم. وقد أشار فؤاد أبو حطب ١٩٧٨ إلى احتمال أنه يوجده فى اللغة العربية ما يشبه ذلك أو يزيد. والنموذج الراهن الذى شغل المؤلف أكثر من عشر سنوات قبل أن يظهر بشكل منظم لأول مرة عام ١٩٧٣ ، هو محاولة لاحداث هذا التكامل المنشود. وقد تعدل النموذج عدة مرات طوال السنوات العشرين الماضية، قدم خلالها فى صورة تقارير تراكمية فى عدة سينمارات ومؤتمرات علمية محلية واقليمية ودولية.

- الافتراضات الأساسية للنموذج:

يستند النموذج الرباعى العملياتى فى صورته الراهنة إلى عدد من الافتراضات الأساسية هى:

١- القدرات العقلية هى فى جوهرها «عمليات معرفية». أما المحتوى الذى تتكرر إليه الإشارة كثيراً فى بحوث القدرات العقلية والذى يشير عادة إلى المضمون الظاهر لمادة الاختبار أو المهمة المعملية أو النشاط العملى للإنسان فما هو إلا القالب الذى

تصب فيه «العملية». وأى افتراض للبنية فى مقابل الوظيفة أو المحتوى فى مقابل العملية هو محض اصطناع لا مبرر له.

٢- الموقف المشكل الذى يستثير السلوك «المعرفى» عند الفرد قد ينشأ عن نقص المدخلات أو الأدلة أو الوسائل أو العادات، ونفضل أن نعبر عن هذه المفاهيم جميعا بمصطلح واحد أكثر شمولاً وهو مفهوم المعلومات الذى يعلب دور متغير التحكم Controlأو المتغير المستقل، ثم يصل المفحوص إلى السلوك النهائى أو الاستجابة أو المخرجات أو الحل بالمعنى الواسع والذى يلعب دور متغير التنفيذ Executionأو المتغير التابع.

٣- العملية المعرفية (وهى جوهر القدرة العقلية) لا يمكن استنتاجها استنتاجاً ملائماً من متغيرات التنفيذ (المتغيرات التابعة) وحدها كما فعل السيكومتريون أصحاب منهج التحليل العاملى التقليدى. أضف إلى ذلك أنه لا يفيدنا كثيراً أيضاً محض العلاقات الوظيفية بين متغيرات التحكم (المتغيرات المستقلة) ومتغيرات التنفيذ (المتغيرات التابعة) على النحو الذى يفترضه المنهج التجريبى التقليدى ومنحى تجهيز المعلومات. فهناك متغيرات أحكام ما قبل التحكم والتى تحدد العمليات المعرفية الكبرى، ومتغيرات أحكام ما بعد التنفيذ أو الحل والتى تعد تقويماً للأداء. وعلى ذلك فالعملية المعرفية (القدرة العقلية) لابد أن تستنتج من أربعة أبعاد للمتغيرات هى أحكام ما قبل التحكم، ومتغيرات التحكم، ومتغيرات التنفيذ، ومتغيرات أحكام ما بعد التنفيذ (التقويم).

- تطور مبادئ تصنيف العمليات المعرفية فى النموذج:

عندما ظهرت الطبعة الأولى من كتاب القدرات العقلية فى عام ١٩٧٣ اقترح فؤاد أبوحطب تصنيف القدرات العقلية كعمليات معرفية فى ضوء بعدين فقط هما متغير المعلومات (أى المتغيرات المستقلة) ومتغير الحلول (أى المتغيرات التابعة).

وظل هذا الاتجاه سائداً حتى (١٩٨٠)وكان التركيز على هذين البعدين مرجعه إلى محاولة تحدى التعريف السائد فى الميدان الذى اعتمد على أساليب الأداء (أى المتغيرات التابعة) وحدها وهو التعريف الذى استندت إليها جميع نماذج التحليل العاملى.

وصنفت متغيرات المعلومات فى ضوء هذا التصور تبعاً لأسس ثلاثة هى: النوع والمقدار والمستوى، كما صنفت متغيرات الحلول تبعاً لأسس ثلاثة أيضاً هى نوع الأداء ونوع الحل وبارامترات القياس، وظلت هذه الأسس ضمن النموذج فى تعديلاته اللاحقة بعد تطويرها والإضافة إليها.

وفى عام ١٩٨٣ أضيف بعد ثالث هو أحكام ما بعد الحل وشملت محكات الدقة والمألوفية والابداعية والواقعية والمرونة. كما اقترح توسيع نطاق النموذج بحيث يشمل نماذج معرفية ثلاثة هى التعلم والتذكر والتفكير، وكانت الفكرة الأخيرة البذرة التى نبت منها البعد الرابع للنموذج.

أصول فكرة الأنواع السبعة للذكاء:

حينما عرض فؤاد أبو حطب نموذجه لأول مرة عام ١٩٧٣ طرح مشكلة العلاقة بين القدرات العقلية كموضوح ينتمى تقليدياً لميدان المعرفة، وسمات الشخصية باعتبارها تنتمى إلى المجال الوجدانى وكان التصور المبدئى يومئذ أن الذكاء هو دالة لنشاط لشخصية ككل، فالمعرفة والوجدان طرفان لمتصل واحد يقع بينهما الذكاء الاجتماعى.

كان التصنيف المبكر اذن لأنواع الذكاء أنها من ثلاثة فئات هى : «الذكاء المعرفى»، و«الذكاء الاجتماعى»، و«الذكاء الوجدانى»، على الرغم مما فى العبارة الأخيرة من تناقص ظاهرى لدى الملتزمين بثنائية «المعرفة» فى مقابل «الوجدان» بدلاً من متصل «المعرفة – الوجدان».

إلا ان هذا التصنيف الثلاثى لأنواع الذكاء لم يستمر (فى الطبعة الثانية من كتاب القدرات العقلية، ١٩٧٨) ففى ضوء متغير نوع المعلومات كمتغير مستقل تم تصنيف أنواع الذكاء إلى سبعة فئات هى:

١- **الذكاء الحسى** Sensory ويتصل بالحساسية من مختلف الأنواع داخل كل وسيط من وسائط الحس، أى أنها ترتبط بأنشطة محتوى متخصصة للحواس المختلفة. والوحدة الأساسية هنا هى العتبة المطلقة. وعلى ذلك فإن الذكاء الحسى قد يكون ذكاء بصرياً أو ذكاء سمعياً، الخ.

٢- **الذكاء الحركى** Motor: ووحداته الرئيسية هى الأفعال المنعكسة التى تبنى عليها الحركات الأساسية (الانتقال المكانى والمعالجة)، وحركات التآزر، والمهارات الحركية العالية (القدرات الجسمية) والمهارات الحركية الدقيقة والاتصال الحركى.

٣- **الذكاء الإدراكى** Perceptual: ويشمل ما يسميه جيلفورد محتوى الأشكال والتى قد تكون بصرية (ووحداتها لها شكل وحجم ولون، الخ)، أو سمعية موسيقية (كاللحن والايقاع وأصوات الكلام) أو لمسية، أو من نوع الاحساس الحركى. والفرق الأساسى بين المعلومات الإدراكية والمعلومات الحسية يكمن فى درجة تعقد المحتوى الذى ينشط فيه كل منهما. فالمعلومات الحسية تكون محض احساس، بينما المعلومات الإدراكية تتمثل فى أبسط مستوياتها فى «صورة الشئ المتميز بذاته». فوحدة محتوى الأشكال البصرية مثلا هى «شكل على أرضية».

٤- **الذكاء الرمزى** Symbolic: ووحداته الأساسية تتمثل فى الحروف الأبجدية أو الأعداد أو المقاطع أو الفونيمات (الأصوات الكلامية) أو التدوين الموسيقى أو غير ذلك من عناصر التشفير.

٥- **الذكاء السيمانتى** Semantic: ووحداته الأساسية تتمثل فى الأفكار والمعانى والتى تتشكل فى أغلب الأحوال فى صورة لغوية، إلا أن بعض المعلومات السيمانتية قد تكون غير لغوية (الصور ذات المعنى).

٦- **الذكاء الشخصى** Personal: هذه الفئة لم تتميز فى تصنيف جيلفورد كما تميزت فئتى الذكاء الرمزى والذكاء السيمانتى مثلاً إنما عدت من نوع المعلومات السيكولوجية كما يسميها، والتى تشتمل على ما يمكن أن يسمى المعلومات الشخصية والمعلومات الاجتماعية، كما حددها فؤاد أبو حطب حينئذ (أى عام ١٩٧٨) أن هاتين الفئتين يجب التمييز بينهما. فالمعلومات الشخصية تتعامل مع ما يتضمن الوعى بالذات. وفى ذلك يقول جيلفورد (Guilford,1967)أننا لا نعرف فقط ولكننا أيضاً نعرف اننا نعرف، كما نعرف أن لدينا مشاعر وانفعالات ومقاصد وأفعال.

والواقع أن اضافة هذه الفئة قدم حلال لمشكلة ثنائية «العقل – الوجدان» التى أشرنا إليها.

٧- **الذكاء الاجتماعى** Social: ويتعامل مع المعلومات التى تتضمن الوعى

بالآخرين، وتشمل ما يسميه سبيرمان العلاقة السيكولوجية، وفى رأيه أن معرفة الآخرين تتم عن طريق المنهج التمثيلى. ويمكن أن تشمل هذه الفئة أيضاً ما يسمى فى علم النفس الاجتماعى المعاصر الإدراك الاجتماعى وإدراك الأشخاص.

وقد ظل هذا التصنيف السباعى لأنواع الذكاء قائماً فى الطبعتين التاليتين من كتاب القدرات العقلية (١٩٨٠، ١٩٨٣)مع التخفف من استخدام كلمة «ذكاء» فقد أشير إليها فى هاتين الطبعتين بالمعلومات الحسية، المعلومات الإدراكية، وهكذا حتى المعلومات الاجتماعية.

ويذكر فؤاد أبوحطب:

آثرنا أن نذكر هذه الحقائق التاريخية حتى يدرك القارئ أن معظم أنماط الذكاء التى أوردها جاردنر فى كتابه أطر العقل عام ١٩٨٣ (Gardner,1983) سبقه المؤلف إليها بسنوات. إلا أننا لم نكن من أنصار استقلال هذه الأنواع السبعة من الذكاء وتوقعنا الحصول على أنواع من الذكاء على درجات متفاوتة من العمومية وصولاً إلى العامل العام.

ومنذ صدور الطبعة الرابعة من كتاب القدرات العقلية عام ١٩٨٣ كان المؤلف مشغولاً بمسألة إعادة التصنيف مع شعور كامن بجدوى التصنيف الثلاثى المبكر. وكانت الخطوة الأولى للعودة إلى هذا التصنيف الثلاثى حين عرض نموذجه فى سمينار هيأته له جامعة لندن عام ١٩٨٣ حين كان أستاذاً زائراً هناك. ثم اتخذ خطوة أكثر تقدما فى البحث الذى عرضه فى الكونجرس الدولى الثالث والعشرين لعلم النفس بالمكسيك عام ١٩٨٤ (Abou-Hatab, 1984).

وفى هذا البحث الأخير استخدم التصنيف الثلاثى الذى لازم النموذج طوال السنوات التالية حتى عام ١٩٩٦ وهو: الذكاء الموضوعى أو غير الشخصى (العلاقات مع الموضوعات المحايدة) Impersonal، والذكاء الاجتماعى (العلاقات بين الأشخاص) Interpersonal، والذكاء الشخصى (العلاقات داخل الشخص الواحد) intrapersnal. فى ضوء نوع المعلومات باعتباره أحد متغيرات التحكم. وقد عرض النموذج بصورته الجديدة المفصلة هذه فى عدة مواضع يمكن للقارئ المهتم الرجوع إليها (فؤاد أبو حطب، ١٩٨٨) (Abou-Hatab,1991,1992).

أبعاد النموذج الرباعى العملياتى:

أشرنا إلى أن النموذج الرباعى العملياتى للمؤلف يتكون من أربعة أبعاد هى: متغيرات الأحكام القبلية (ما قبل التحكم) أو متغيرات القرار، ومتغيرات التحكم (المعلومات)، ومتغيرات التنفيذ (الاستجابة أو الحل). ومتغيرات الأحكام البعدية (ما بعد التنفيذ) أو متغيرات التقويم. وفيما يلى عرض لهذا الأبعاد الأربعة:

البعد الأول: متغيرات الأحكام القبلية أو متغيرات القرار:-

تحدد متغيرات ما قبل التحكم Pre-control النموذج الفرعى للعمليات المعرفية الكبرى التى يقع فى نطاقها مجال البحث أو الدراسة. وفى هذا علينا أن نقرر ونختار بين النموذج الفرعى للتفكير Thinking Submodel وبعض العمليات التى تشبهه مثل الأحساس والانتباه والإدراك، والنموذج الفرعى للتعلم Learning Sub-model والنموذج الفرعى للذاكرة Memory Submodel.

وتتشابه النماذج الفرعية الثلاثة فى أسس تصنيف العمليات المعرفية والمتضمنة فى كل منها كما تتحدد بالأبعاد الثلاثة التالية (متغيرات التحكم ومتغيرات التنفيذ ومتغيرات أحكام ما بعد التنفيذ أو التقويم) إلا أن نواتج كل منها مختلفة. فالنواتج فى الإحساس والانتباه والإدراك والتفكير يمكن ان نسميها استراتيجيات Strategies أو أساليب Styles، وفى التعلم مهارات Skills، وفى الذاكرة كفاءات Com-petencies، ولعلنا بهذا نميز بين المفاهيم التى شاعت فى تراث علم النفس المعرفى المعاصر على نحو يدعو إلى الخلط وعدم الوضوح. وهى جميعاً تنتمى لعمليات تجهيز معلومات Information Processing.

ونستخدم فى التمييز بين النماذج الفرعية الثلاثة محكاً يعتمد فى جوهره على متصل الجدة - المألوفية للمعلومات كما يلى:-

١- اذا كانت الفجوة المعلوماتية Informational Gap التى تؤلف المشكلة جديدة أو مفاجئة (ويشمل ذلك المحاولة الأولى فى بحوث التعلم أو العرض الواحد القصير جداً فى بحوث الذاكرة) فان اهتمامنا فى هذه الحالة ينتمى إلى النموذج الفرعى الأول. وقد أطلقنا على نواتج هذا النموذج «استراتيجيات أو أساليب معرفية»، ويقع فى نطاق هذا النموذج الفرعى العمليات المعرفية الكبرى للاحساس، الانتباه ، الإدراك،

التفكير والتى أوضحنا الفروق الجوهرية بينها فى كتابنا القدرات العقلية (٣٣).

٢- إذا تكرر عرض المعلومات (محاولات أو عروض) بحيث تتزايد المألوفية بالفجوة المعلوماتية (المشكلة) فان النموذج الفرعى السائد عندئذ هو النموذج الثانى وهو فى جوهره يتناول العملية المعرفية الكبرى للتعلم. وفى هذا الصدد نحب أن ننبه إلى أنه لا يوجد تعلم من محاولة واحدة أو عرض واحد ولو كان قصيراً. فالمحاولة الأولى أو العرض الأول تنتمى إلى النموذج السابق وخاصة حين تكون المعلومات فيها على درجة كافية من الجدة. وتزداد المألوفية بزيادة عدد مرات العرض أو المحاولة. وكل التراث السيكولوجى الذى كتب حول التعلم من محاولة واحدة أو عرض واحد قصير يمكن تصنيفه فى ضوء النموذج الراهن فى نموذج التفكير أو الإدراك أو الانتباه أو الاحساس. والاستراتيجيات قد تتحول إلى مهارات بالتدريب، وكل منها يوصف بالكفاءة إذا اجتاز امتحان الزمن من خلال الذاكرة. وبالطبع فإن مهارات التعلم قد تعتمد على استراتيجيات الاحساس أو الانتباه أو الإدراك أو التفكير السابقة عليها وتؤثر فيها، وكلاهما قد يحدد كفاءات الذاكرة التى تتلوها.

٣- إذا كانت المعلومات سبق عرضها وتم تخزينها Storage بحيث تتلاشى الفجوة المعلوماتية الأصلية، ويكون المطلوب من المفحوص فى فجوة معلوماتية جديدة استرجاعها أو استردادها Retreival، فإننا نكون بإزاء النموذج الفرعى الثالث لعملية معرفية كبرى هى الذاكرة. ولا يمكن للذاكرة أو تنشط فى فراغ معلوماتى، وإنما عادة ما تدرس فى معمل علم النفس أو فى المواقف الطبيعية بعد اتمام ما نسميه التعلم الأساس Basic Learning أو الاكتساب Acquisition حين يتطلب الأمر إعادة قياسه بعد فترة من الزمن. وبالطبع فإن العرض المتتابع فى مواقف التعلم يتضمن قدراً من الذاكرة من المحاولات السابقة إلى المحاولة اللاحقة نتيجة المألوفية المتزايدة. ومعنى ذلك أن المعلومات فى مواقف التذكر تكون من النوع الذى سبق التعرض له وتتوافر لها المألوفية الكاملة.

وبهذا تصبح الذاكرة عملية تخزين المعلومات واسترجاع أو استعادة هذه المعلومات المختزنة. ولعل هذا هو سبب تسمية نواتج النموذج الفرعى للذاكرة بالكفاءات Competencies والتى تتطلب لنجاحها وجود رصيد مختزن من آثار

الخبرة السابقة التى تتجاوز اختبار الزمن سواء كان قصيراً (ذاكرة المدى القصير) أو طويلاً (ذاكرة المدى الطويل). وتعتمد بالطبع على استراتيجيات الإحساس أو الانتباه أو الإدراك أو التفكير ومهارات التعلم السابقة عليها كما تؤثر فيها.

البعد الثانى: متغيرات المعلومات (التحكم) :-

متغيرات التحكم هى ما اعتدنا أن يشار لها فى تناول النموذج باسم المتغيرات المستقلة أو متغيرات المعلومات ويستخدم مفهوم المعلومات هنا بنفس المعنى الذى استخدمه فيه جيلفورد، أى «ما يستطيع الانسان تمييزه». وهذا يعنى أن التسجيل والتشفير هى متطلبات أساسية للنموذج الحالى.

ومهمة متغيرات المعلومات (أو التحكم) فى النموذج الحالى هى إحداث الفجوة المعلوماتية (أو المشكلة)، وتصنف تبعاً للمبادئ الآتية :-

١- نوع المعلومات :-

يرتبط هذا المبدأ بالمبدأ الشهير للتصنيف المستخدم فى مجال القدرات العقلية وهو مبدأ المحتوى أو المضمون. وقد اقترحت فئات ثلاثة عند ثورنديك (العملى والمجرد والاجتماعى)، وعند ثرستون وأيزنك والقوصى (الأعداد والأشكال والكلمات) وعند جيلفورد (الأشكال البصرية والسمعية والرموز والمعانى والسلوك الاجتماعى).

وفى النموذج الحالى نفضل تصنيف أنواع المعلومات إلى الفئات الآتية :-

(أ) المعلومات الموضوعية أو غير الشخصية Impersonal (الذكاء الموضوعى): وتشمل الأشياء والرموز (ومنها الرموز اللغوية) وجميع المواد التى يستخدم معها المفحوص عملية الفحص الخارجى Extrospection ويتعامل معها كموضوعات أو أشياء أو رموز لهذه الموضوعات أو الأشياء أو معان تدل عليها. وهذا النوع من المعلومات يشمل كل ما تناوله مجال الذكاء بالمعنى التقليدى، ويمتد على متصل من المعلومات الحسية الإدراكية إلى المعلومات الرمزية اللغوية. وتسمى نواتج النماذج الفرعية الثلاثة فى هذه الحالة بالاستراتيجيات الموضوعية والمهارات الموضوعية والكفاءات الموضوعية.

(ب) المعلومات الاجتماعية Socialأو التى تدل على العلاقات بين الأشخاص Interpersonal: وتتضمن ما يسمى تقليديا الذكاء الاجتماعى والذى يشمل الادراك الاجتماعى وإدراك الأشخاص وكل المواد التى تستخدم اجتماعياً والتى يتم التعامل معها بطريقة الفحص المتبادل Interspectionأو الفحص بالمشاركة. وتسمى نواتج النماذج الفرعية الثلاثة فى هذه الحالة بالاستراتيجيات الاجتماعية والمهارات الاجتماعية والكفاءات الاجتماعية.

(ج) المعلومات الشخصية Personal أو المعلومات داخل الشخص الواحد Intrapersonal: وهى التى تتضمن ما سمى منذ عام ١٩٧٣ الذكاء الوجدانى أو الشخصى. ولم يظهر هذا المصطلح فى التراث السيكولوجى الغربى فى الولايات المتحدة إلا عام ١٩٨٣ فى كتاب (frames Of mind)لمؤلفه هوارد جاردنر، ولا ندرى أن كان هذا المؤلف قد تأثر بكتابات فؤاد أبو حطب، فهو لم يشر إلى ذلك مطلقاً، أم أنه وصل إلى نفس الفكرة مستقلا بعد ظهور الكتاب باللغة العربية طوال هذه السنوات. وعموماً فإن المهام التى تنتمى إلى هذه الفئة هى ما يستخدم المفحوص فيه طريقة الفحص الداخلى Introdpection(الاستبطان) . وتسمى النماذج الفرعية الثلاثة فى هذه الحالة بالاستراتيجيات الشخصية والمهارات الشخصية والكفاءات الشخصية.

٢- مستوى المعلومات:

يشير هذا المبدأ إلى مظهر التعقد والبساطة فى المعلومات. وهو متضمن فى نموذج القوصى فيما يسميه الهيئة Form، وفى نموذج جيلفورد فيما يسميه الناتج Product . ويقتصر هذا النموذج على الفئات الأربع الآتية للمستوى من بين نواتج جيلفورد لأنها تتضمن بوضوح فكرة الترتيب الهرمى التراتبى أو الهيراركى وهى:-

(أ) الوحدات: وهى أبسط ما يمكن أن تحلل إليه المعلومات المعطاة أو المولدة أو المستخرجة.

(ب) الفئات: وهى مجموعات من الوحدات تجمعها خصائص مشتركة.

(ج) العلاقات: وهى الروابط التى تربط بين الوحدات أو الفئات تبعاً لمبدأ معين مثل الهوية أو التشابه أو التضاد أو غيرها.

(د) المنظومات: وهى مركبات تجمع وحدات أو أجزاء متفاعلة أو بينها علاقات متداخلة. وقد يتألف النسق من مركب من الفئات حين تكون الفئات معقدة نظراً لتعدد الخصائص المشتركة ويسمى فى هذه الحالة البنية Srructure، كما قد يتألف من مركب من العلاقات تبعاً لوجود أكثر من علاقة واحدة وهو المعنى الأكثر شيوعاً لمفهوم المنظومة أو النسق.

٣- طريقة العرض:-

يشير هذا المبدأ إلى نظام عرض المعلومات، وفى هذا يتم التمييز بين فئتين:

(أ) العرض التكيفى Adaptiveأو المنتظم Systematic:-

وفيه تقدم للمفحوص تعليمات صريحة حول طبيعة المهمة أو العمل أو طرق التعامل مع المعلومات.

(ب) العرض التلقائى Spontaneousأو العشوائى Random:

وفيها لا يقدم إلا القليل من المعلومات حول طبيعة المهمة ويترك للمفحوص تحديد طبيعتها.

٤- مقدار المعلومات:-

هذا المبدأ مبدأ كمى اذا قارناه بالمبادئ الثلاثة السابقة (فهى فى طبيعتها كيفية). وقد كان هذا المبدأ منذ البداية تحدياً للنموذج الحالى، كما يحتل منزلة خاصة لما يحيط به من مشكلات تتصل بقياسه. وقد نشأت هذه الصعوبات من حقيقة مؤداها أن بنية المشكلة التى تحتاج إلى الحل ترتبط ارتباطاً وثيقاً بنمط الحل وخاصة التمييز بين الحل التقاربى والحل التباعدى وقد بذل صاحب النموذج جهداً فى اقتراح الطرق الملائمة لقياس مقدار المعلومات لكل من نوعى الحل.

ويجب أن ننبه إلى أن مفهوم مقدار المعلومات المستخدم فى هذا السياق يختلف عن ذلك المستخدم فى النظرية الكلاسيكية للمعلومات وتطبيقاتها فى علم النفس. إنه يشير إلى مقدار (أو عدد) الوحدات والفئات أو العلاقات أو المنظومات (وكلها مستويات) للمعلومات الموضوعية والاجتماعية والشخصية المستخدمة فى التحكم فى النسق المعرفى للإنسان.

وتستخدم فى هذا النموذج عدة طرق لقياس المعلومات هى:

١- مـقـدار طلب المعلومـات Amount of Information Demand (AID)مما يتمثل فى العرض التتابعى للمعلومات تبعاً لطلب المفحوص لها وحاجته إليها، وعندما تتوقف هذه الحاجة يتوقف العرض.

٢- مقدار ثروة (أو رصيد) المعلومات Amount of Information Sup-ply (AIS)وتقاس هذه الثروة أو الرصيد أو الذخيرة بمعادلة رياضية تقدر مقدار تدفق هذه الثروة فى فترات زمنية متساوية ينقسم إليها المدى الكلى لزمن الأداء.

٣- درجة التأهب فى تعليمات الاختبار أو المهمة.

البعد الثالث: متغيرات التنفيذ (الحلول أو الاستجابات):

يشير هذا البعد إلى طرق حل المشكلة، أو طرق سد الفجوة المعلوماتية.

وتصنف متغيرات التنفيذ (المتغيرات التابعة أو متغيرات الاستجابات أو الحلول) فى ضوء المبادئ الآتية:-

١- طريقة التعبير عن الحل (نوع الأداء):

نميـز فى هذا النمـوذج بين الأداء الحـركى Psychomotorوالأداء اللفظى Verbal. وهو تمييز تقليدى يشبه ما كان يسمى الأداء اللفظى وغير اللفظى. وإدخال المجال الحركى هنا يشير إلى الاختلاف الجوهرى مع بنجامين بلوم الذى اعتبر المجال الحركى موازيا للمجالين المعرفى والوجدانى. ويمكن أن يضاف بالطبع الأداء الفسيولوجى Physiologicalاذا استخدمنا الرسام الكهربائى للمخ أو مقياس الاستجابة السيكوجلفانية أو غيرهما من المقاييس الفسيولوجية. وقد ظهر الاهتمام فى السنوات الأخيرة بنوع رابع من الأداء يستخدم فى التحليل الكيفى لحلول المشكلات هو تحليل البـروتوكـولات Protocol Analysisبعـد الحل، أو التلفظ به أثناء الحل Verbalization.

٢- نوع الحل:-

يميز النموذج بين نوعين من الحلول هما:-

التقويم النفسى ـــــــ

البعد الرابع: المتغيرات البعدية أو متغيرات التقويم (أحكام ما بعد التنفيذ):

توجد مجموعة من المتغيرات تشمل الأحكام التى يصدرها المفحوص على أدائه أو حله، أو يصدرها الآخرون على هذه الحلول او الأداءات، أو يصدرها هو على أداءات وحلول الآخرين (التى تقدم له عند عرض المشكلة).

وتصنف متغيرات هذه الفئة إلى ما يلى:-

١- محل (موضع) الحكم:

ويصنف إلى فئتين حسب محل الحكم Locus of Judgment فقد يكون هذا المحل داخلياً حين يحكم المفحوص على أدائه، وقد يكون خارجياً حين يحكم الآخرون على أداء المفحوص، أو يحكم هو على أداء الآخرين.

٢- نوع المحك:

وتشمل هذه الفئة المحكات المختلفة التى تستخدم فى الحكم أو التقويم.

وتوجد فئات كثيرة من المحكات منها:

(أ) الصواب فى مقابل الخطأ.

(ب) الندرة فى مقابل الشيوع.

(ج) التنوع فى مقابل التجانس.

٣- مستوى الحكم:-

تشير هذه الفئة إلى المستوى الذى يكون عليه تقويم الفرد لأدائه أو تقويم الآخرين له. ويوجد نوعان من المستوى:

(أ) الأحكام الكيفية المنفصلة والتى يصنف فيها الأداء إلى فئات مستقلة من الجودة.

(ب) الأحكام الكمية المتصلة والتى تكون فى صورة درجة جودة فى ضوء المحك المستخدم.

٤- درجة الثقة فى الحكم:

تعتبر هذه الفئة من فئات متغيرات الأحكام البعدية من نوع المتغيرات الكمية، وترتبط بدرجة ثقة أو يقين المفحوص فى الحكم الذى يتوصل إليه هو، أو فى الأحكام التى يتوصل إليها الآخرون.

التقويم النفسى ـــــــ

وهناك العديد من الدراسات التى أجريت فى إطار هذا النموذج سواء قام بها فؤاد أبو حطب مستقلا أو بالاشتراك مع طلابه وقد نشرت هذه الدراسات فى مواقع مختلفة يمكن الرجوع لها سواء تلك التى نشرت ف مصر او فى المؤتمرات الدولية والاستفادة من هذا النموذج فى بناء المقاييس المصرية بدلاً من الاعتماد على نقل وترجمة المقاييس الأجنبية.

مراجع الباب الثالث

١ – أحمد زكى صالح: الأسس النفسية للتعليم الثانوى، دار النهضة العربية، ١٩٧٢ .

٢ – أحمد زكى صالح: علم النفس التربوى، مكتبة النهضة المصرية، ١٩٧٢ .

٣ – أحمد زكى صالح: علم النفس فى الإدارة والصناعة، مكتبة النهضة المصرية، ١٩٧٢ .

٤ – أحمد زكى صالح: تعليمات اختبار الذكاء المصور، (ب. ت) .

٥ – أحمد فكرى: الطريقة المبتكرة لقياس العقول، ١٩٢٠ .

٦ – إسماعيل القبانى: مقياس ستانفورد ـ بينيه للذكاء، لجنة التأليف والترجمة والنشر، ١٩٣٨ .

٧ – إسماعيل القبانى: قياس الذكاء، المطبعة الأميرية، ١٩٣٨ .

٨ – إسماعيل القبانى: اختبار الذكاء الابتدائى، لجنة التأليف والترجمة والنشر (ط ثانية)، ١٩٣٩ .

٩ – السيد محمد خيرى: تعليمات اختبار الذكاء الإعدادى، دار النهضة العربية (ب، ت) .

١٠ – السيد محمد خيرى: تعليمات اختبار الذكاء العالى، دار النهضة العربية (ب، ت) .

١١ – المقتطف: مقاييس الذكاء، أغسطس ١٩١٧ .

١٢ – آمال أحمد مختار صادق، الصورة السعودية لمقياس ستانفورد ـ بينيه (بحث غير منشور)، ١٩٧٧ .

١٣ – ــــــــــــ ، دراسة مقارنة لاختبار رسم الرجل فى كل من الزى التقليدى والأوربى، فى بحوث فى تقنين الاختبارات النفسية (تحرير فؤاد أبو حطب)، المجلد الثانى، مكتبة الأنجلو المصرية، ١٩٧٩، ١٠٥ – ١٢٨ .

١٤ – آمال أحمد مختار صادق: اختبارات كاتل للذكاء العام – مقياس الذكاء المتحرر من أثر الثقافة – المقياس الثانى والثالث – صورة أ، ب –كراسة التعليمات – الانجلو المصرية ٢٠٠٧ .

١٥ – آمال جلال محمد : تقنين اختبار المصفوفات المتتابعة الملون على تلاميذ الحلقة الأولى من التعليم الاساسى من المستويات الثقافية المختلفة – رسالة ماچستير

غير منشورة – كلية التربية – جامعة حلوان، ٢٠٠٠ .
١٦– أمير بقطر: اختبارات الذكاء، مجلة التربية الحديثة، ١٩٢٨ .
١٧– أمين محمد صبرى نور الدين: بعض الخصائص السيكومترية لمقياس ستانفورد – بينيه الرابع لدى عينة من الأطفال فى مرحلة ما قبل المدرسة – رسالة ماجستير غير منشورة – كلية التربية جامعة عين شمس، ١٩٩٥ .
١٨– أمين محمد صبرى نور الدين: فاعلية استخدام الاختبار الموائم باستخدام الحاسب فى تقدير قدرة الافراد وتحديد الخصائص السيكومترية للمقياس ، رسالة دكتوراه – كلية التربية جامعة عين شمس، ٢٠٠٢ .
١٩– أمينة محمد كاظم : المقاييس البريطانية للقدرات – المركز القومى للإمتحانات والتقويم التربوى – مؤتمر العمل المصغر لإعداد مقياس قومى للذكاء، ١٩٩٥ .
٢٠– أيمن الديب محمد شوشة: استخدام نموذج Pass فى التشخيص الفارقى لعينة من ذوى الاحتياجات الخاصة المعاقين ذهنيا القابلين للتعلم. رسالة ماجستير غير منشورة – كلية البنات جامعة عين شمس، ٢٠٠١ .
٢١– أيمن الديب محمد شوشه وصفاء يوسف الأعسر: دليل منظومة التقييم المعرفى م.ت.م (الكاس) CAS للذكاء – الأنجلو المصرية ٢٠٠٦ .
٢٢– رشا محمد عبد الله: القيمة التشخيصية لنموذج التقدير المعرفى CAS كما تظهر فى العلاقة بين نموذج PASS والتحصيل الدراسى رسالة ماجستير غير منشورة – كلية البنات جامعة عين شمس ٢٠٠٣ .
٢٣ – جابر عبد الحميد جابر، يوسف الشيخ: علم النفس الصناعى، دار النهضة العربية، ١٩٦٨ .
٢٤ – جابر عبد الحميد جابر: الذكاء ومقاييسه، دار النهضة العربية، ١٩٧١ .
٢٥ – جمال محمد على: تقنين مقياس ستانفورد ـ بينيه للذكاء على تلاميذ مرحلة التعليم الأساسى، رسالة ماجستير، كلية التربية جامعة عين شمس، ١٩٨٣ .
٢٦ – حسن حسين زيتون: اختبار مراحل بياجيه للنمو العقلى، الكويت: دار القلم، ١٩٨٦ .
٢٧ – حسن عمر: مقياس الذكاء، ١٩٢٨ .
٢٨ – حسنين الكامل: اختبار هايدلبرج للذكاء، مكتبة النهضة المصرية، ١٩٨٥ .

التقويم النفسى ـــــــ

(ط ٤) ١٩٩٨، (ط ٥) ٢٠٠٨ .

٤٢ – فؤاد البهى السيد: الذكاء، دارالفكر العربى، ١٩٧١ .

٤٣– فتحى السيد عبدالرحيم: اختبار رسم الرجل، تطبيقه وتقنينه على البيئة اليمنية، بحث غير منشور ، ١٩٧٥ .

٤٤– فرج عبد القادر طه: العلاقة بين الإصابات فى الصناعة وكل من الصفحة النفسية للذكاء والسرعة الإدراكية والسرعة الحركية، المجلة الاجتماعية القومية ـ المجلد السادس، سبتمبر ١٩٦٩، ٧٣ – ١٠٤ .

٤٥ – ـــــــ ، صلاح أحمد مرحاب، الصورة المغربية لمقياس وكسلر ـ بلفيو لذكاء الراشدين والمراهقين، المغرب– الرباط، مطبعة الكوثر، ١٩٧٧ .

٤٦ – فوس، برايان (ترجمة فؤاد أبو حطب): آفاق جديدة فى علم النفس ، عالم الكتب، ١٩٧٢ .

٤٧ – لويس كامل مليكة: الدلالات الإكلينيكية لمقياس وكسلر ـ بلفيو لذكاء الراشدين والمراهقين، النهضة المصرية، ١٩٦٠ .

٤٨ – لويس كامل مليكة: المقررات فى قياس الذكاء، النهضة المصرية، ١٩٦٠ .

٤٩ – لويس كامل مليكة: نماذج التصحيح وجداول الدرجات الموزونة ونسب الذكاء لمقياس وكسلر ـ بلفيو لذكاء الراشدين والمراهقين، النهضة المصرية، ١٩٦٠ .

٥٠ – مائسة أنور المفتى: إعداد صورة معدلة لاختبار وكسلر ـ بلفيو لذكاء الراشدين والمراهقين، رسالة ماجستير، كلية الآداب جامعة عين شمس، ١٩٧١ .

٥١ – مالك البدرى: سيكولوجية رسوم الأطفال: اختبارات رسم الإنسان وتطبيقاتها على أطفال البلاد العربية، بيروت: دار الفتح للطباعة والنشر، ١٩٦٦ .

٥٢ – ماهر محمود الهوارى: دراسة لصدق مقياس ستانفورد ـ بينيه للذكاء، رسالة ماجستير غير منشورة. كلية الآداب جامعة عين شمس، ١٩٦٣ .

٥٣ – محمد خليفة بركات: الاختبارات والمقاييس العقلية، مكتبة مصر، ١٩٥٤ .

٥٤ – محمد عبدالسلام أحمد، لويس كامل مليكة: مقياس ستانفورد ـ بينيه للذكاء، مكتبة النهضة المصرية، ١٩٥٦ .

٥٥ – محمد عماد الدين إسماعيل، لويس كامل مليكة: مقياس وكسلر ـ بلفيو لذكاء الراشدين والمراهقين، مكتبة النهضة المصرية، ١٩٥٦ .

٥٦ – محمد عماد الدين إسماعيل، لويس كامل مليكة: مقياس وكسلر- بلفيو لذكاء الأطفال، مكتبة النهضة المصرية، ١٩٦١ .

٥٧ – محمد عماد الدين إسماعيل، لويس كامل مليكة: بحوث فى اختبارات وكسلر لذكاء الراشدين والمراهقين، مكتبة النهضة المصرية (ب، ت) .

٥٨ – محمد متولى أحمد غنيمة: تقنين اختبار رسم الرجل بالنسبة لتلاميذ المرحلة الابتدائية فى مصر، رسالة ماجستير، كلية التربية جامعة عين شمس، ١٩٧٦ .

٥٩ – محمد نسيم رأفت: تقنين اختبار رسم الرجل على البيئة الكويتية، الكويت، التقرير السنوى لكتب الخدمة الاجتماعية، إدارة التربية للطباعة وزارة التربية بالكويت، ١٩٦٨ .

٦٠ – مصطفى فهمى: اختبار رسم الرجل، دار مصر للطباعة، (ب، ت) .

٦١– ميشيل اسكندر: رسوم الأطفال وعلاقتها بنموهم العقلى، رسالة ماجستير، المعهد العالى للتربية للمعلمين (كلية التربية جامعة عين شمس)، ١٩٣٥ .

٦٢– نادية السيد الحسينى: البنية العاملية لاختبارات الذكاء غير اللفظية فى مرحلتى الطفولة والمراهقة. رسالة دكتوراه غير منشورة، كلية الدراسات الانسانية – جامعة الازهر، ١٩٨٧ .

٦٣– نادية محمود شريف: أهم الاتجاهات فى تفسير الذكاء فى مشروع المقاييس المصرية للقدرات . مؤتمر العمل المصفر لاعداد مقياس مصرى للذكاء – المركز القومى للامتحانات والتقويم التربوى، ١٩٩٥ .

٦٤ – نعيم عطية: ذكاء الأطفال من خلال الرسوم، نسق جديد لاختبار رسم الرجل، دراسة استطلاعية، بيروت: دار الليعة، ١٩٨٢ .

٦٥ – هاريس ، د. ب (ترجمة محمد فرغلى فراج، عبد الحليم محمود السيد، صفية مجدى)، اختبار الرسم لجودانف وهاريس (غير منشور)، ١٩٧٦ .

٦٦– هشام فتحى محمد جاد الرب: تطوير اختبار كاتل للذكاء باستخدام نماذج السمات الكامنة وأثر ذلك على قدرة الاختبار بالتنبؤ بالتحصيل الدراسى. رسالة ماجستير غير منشورة كلية التربية – جامعة المنصورة، ١٩٩٩ .

67. Anastasi, A. Psychological Testing, Macmillan, 1961, 1984, 1982, 1976, 1984 and Urbina ., (7th ed) 1997 .

68. Binet, A., Simon, T. The Development of Intelligence in Chlidren, In; Jenkins, J. J. and Patterson, D. G. (ed.) Studies In Individual Differences. London: Methuen, 1961.

69. Bond, E.A. Tenth Grade Abilities and Achievements. Teacher College, I940.

70. Bradway, K.P., Thompson, C.W., Grays, R.B. Preschool IQ's after Twenty-Five Years. J. Educ., Psychol., 1958, 49, 278-281.

71. Burt, C. Mental and Scholastic Tests. (4th ed.) London, Steples Press, 1962.

72. Carlton, J.S. and Jensen, C.M. The Factorial Structure of the Raven Coloured Progressive Matrices. Educ. and Psychol. Measurement, 1980, 40, 111-115.

73. Cattel, R.B., Theory of Fluid and Crystalized Intellingence Psychology, 54, 1-22, 1963 .

74. Cohen, J.A. Factor-Analytically Based Rational for the Wechsler Adult Intelligence Scale. J. Consul. Psychol. 1957, 2I, 451, 457.

75. Cohen, J. The factorial structure of the WAIS Between Early Adulthood and Old Age. J. Consul. Psychol., 1957, 283-390.

76. Cronbach, L.J. Essentials of Psychological Testing. New York, Harper, Brothers, 1960, 1970, 1984.

77. Edwards, A.J. Individual Mental Testing, vol. I, Scraton . Inext Educ. Publ, 1971.

78. _____ Individual Mental Testing, vol. II, Scraton. In: text Educ. PubL, 1972.

79. Elliot, C.D., Differential Ability Scales: Adminstration and Scoring Manual. San Antonio Psychol. Cor. 1990. In Anastasi & Urbina 1997.

80. Fagen, J.E., Intelligence. A Theoretical Viewpoint, Current Direction in Psychological Science. I, 82-86, 1992 .

81. Fagen, L.E., and Detterman, The Fagen Test of Infant Intelligence. A Theoretical Summery. J. Applied Developmental psychology, 13, 173-193, 1992.

82. Freeman, F.S. Theory and Practice of Psychological Testing. (3rd. ed.), New Delhi Oxford, 1972.

83. Guertin, W.H., et al. Research with the Wechsler Scales for Adults: 1960-1965 Psychol. Bull., 1966, 66, 385-409.

84. _____ etal. Research with Wechsler Scales for Adults: 1965-1970 Psychol. Res., 1971, 21, 289-339.

85. Guilford, J.P. The Nature of Human Intelligence. McGraw- Hill, 1967.

86. Himelstein, P. Research with the Stanford-Binet, Form L.M. Thc First Five Years. Psychol. Bull., 65, 156-164, 1966.

87. Horn, J.L. Theory of Fluid and Crystallized Intelligence. In R.J. Sternberg (ed)., Encyclopedia of Human Intelligence (pp. 443-451). New York: Macmillan 1994.

88. Jones, L.V. A factor Analysis of the Stanford-Binet at four Age level. Psychometrika, 1949, 14, 299-331.

89. Jones, L.V. Primary abilities in the Stanford- Binet, age 13, J. Gent. Psychol., 1954, 74, 126-147.

90. Kennedy, W.A., et al. The Ceiling of the New Stanford-Binet. J. Clin. Psychol., 1960, 17, 274-276.

91. Matarazzo, J.D. Wechsler's Measurement and Appraisal of Adult Intelligence (5th ed.) Baltimore: Williams and Witkins, 1972.

92. McNemar, Q. The Revision of the Stanford-Binet Scale : An analysis of the Standardization data. Boston : Houghton Mifflin, 1942.

93. Moriatory A.E. Children's Way of Coping With the Intelligence Test. Menninger Clinic Bull., 1960, 24, 115-127.

94. Macusker, P.J., Validation of kaufman Ishikuma, and kaufman - Packer's Wechsler Adult Intelligence scale - Revised short Forms on a Clinical Sample. Psychological Assessment 6, 246- 1994, in Anastasi & Urbina 1997 .

95. Minton, B. & Pratt, S. Identification Discrepancies Gifted and Highly Gifted Students: How Do They Score on the SB5? Roeper Review, 28, 232-236 .

96. Mleko, A. & Burns, T. Test Review: Stanford-Binet Intelligence scales (5th ed.) Applied Neuropsychology, 12, 179-180.

97. Peterson, J. Early Conceptions and Tests of Intelligence. New York, Yonkers, World Book, 1926.

98. Pineau, S.R. Changes in Intelligence qotient from Infancy to Maturity. Boston : Houghton Mifflin, 1961.

99. Roid, G. H. Stanford - Binet Intelligence Scales, Fifth Edition: Itasca, IL: Riverside Publishing, (2003-a).

100. Roid, G.H. Stanford - Binet Intelligence Scales, Fifth Edition: Technical Manual. Itasca, IL: Rverside Publishing, 2003 b.

101. Roid, G. H., & Carson, A.D. Special Composite Scores for the SB5. (Stanford- Binet Intelligence Scales, Fifth Edition Assessment Service Bulletin No. 4). Itasca, IL: Riverside Publishing, (2003).

102. Ruf, D.L. Use of the SB5 in the Assessment of High Abilities (Stanford- Binet Intelligence Scales, Fifth Edition Assessment Service Bulletin No. 3). Itasca, IL: Riverside Publishing, (2003).

I03. Raven, J. The Standard Progressive Matrices Test. London Lewis, 1958.

104. Raven J. The Coloured Progressive Matrices Test. London: Lewis,1965 .

105. Raven, J., 1965 J.C., Manual for Raven's Progressive Matrices Series. I and II. J. C. Raven Ltd., London: H. K. Lewis. 1983 .

I06. Sadek, A.A., Factor Analytic study of Musical Abilities of Egyptian students Taking Music as a special subject. Ph.D. Thesis, London university 1968 .

107. Sattler, J.M Assessment of Chilfren, WISC III and WPPSI- R Supplement san Diego, C.A. 1992 .

108. Sattler, J.M. Assessment of Children's Intelligence and special abilities (2nd ed.), Boston : Allyn and Bacon, 1972.

109. Saunders, D.R. On the Dimentionality of the WAIS Battery for two Groups of Normal Males. Psychol. Rep., 1959,(529-541).

110. Saunders, D.R. A Factor Analysis of the Information and Arith-

metic items of the WAIS, Psychol. Rep., 1960, 6, 367- 383.

111. Saunders, D.R. A Factor Analysis of the Picture Completion Items of the WAIS. J. Clin. Psychol, 1960, 16, 146-149.

112. Saunders, D.R. Alpha frequency : A Cross-Validation.J. clin. Psycho.,., 1961, 17, 165-167.

113. Terman, L.M., Merrill, M.A. Measurement of Intelligence, Boston, Hougton Mifflin, 1937.

114. Terman, L.M., Merrill, M.A. Stanford-Binet Intelligence Scale. Manual for the third revision. Form L-M. Boston : Hougton Mifflin, 1960.

115. Terman, L.M. and Merrill, M.A. Stanford-Binet Intelligence Scale : 1972, Norms Edition, Boston : Hougton Mifflin, 1973.

116. Thorndike, R.L., Hagen, E.P. and Sattler, J.M. Stanford- Binet Intelligence Scale, (4th ed.) Chicago : The Riverside Publishing- Co., 1985.

117. Thorndike, R.L., Hagen, E.P., Sattler, J. M., Technical Manual, the Stanford Binet Intelligence Scale, 4th ed. 1986 .

118. Vernon, P.E. Intelligence and Attainment tests. University of London Press, 1960.

119. Wechsler, . The measurement and appraisal of adult intelligence. (4th ed.) Baltimore : Williams and Witkins, 1958.

120. Wolf, T.H. Alfred Binet. Chicago : Univ. of Chicago Press, 1973.

الباب الرابع
مقاييس الاستعدادات

الفصل التاسع

بطاريات الاستعدادات المتعددة

يقصد ببطاريات الاستعدادات المتعددة multiple-aptitude batteries مجموعة المقاييس التى تهدف إلى قياس مجموعة من مظاهر النشاط العقلى، كل على حدة، قياسا مستقلا، وتعطى لكل منها درجة مستقلة، وعادة ما لا تعطى هذه البطاريات مؤشرا واحدا للبطارية ككل (مثل نسبة الذكاء أو العمر العقلى أو الدرجة المعيارية). وفى هذا تختلف بطاريات الاستعدادات المتعددة عن اختبارات الذكاء العام التى تناولناها فى الباب السابق، كما تختلف عنها أيضاً فى أن اختبارات الذكاء العام تعتمد فى بنائها على درجة تجانس أسئلتها أو اختباراتها الفرعية بحيث تعطى فى النهاية مقياسا للذكاء يتوافر فيه شرط الاتساق الداخلى، حتى ولو كانت هذه الاختبارات تعطى درجات منفصلة للأقسام اللفظية أو العملية فيها (مثل مقاييس وكسلر) أو للأقسام اللفظية والعددية (مثل اختبارات ذكاء تلاميذ المرحلة الثانوية وطلاب الجامعات)، بالإضافة إلى الدرجة الكلية أو التقدير العام للذكاء أما فى بطاريات الاستعدادات المتعددة فإن الهدف من القياس هو إعطاء المفحوص عددا من الدرجات المنفصلة للتمييز بين القدرات العددية التى تظهر اختلافات داخل الفرد الواحد بقدر ما تظهرها بين الأفراد.

ويمكن أن نلخص أهم العوامل التى أدت إلى زيادة الاهتمام بهذا الاتجاه «الفارق» وخاصة منذ نهاية الحرب العالمية الثانية فيما يلى:

١ – زيادة الاعتراف بالفروق الفردية داخل الفرد intraindivual variations فى أدائه على الأجزاء المختلفة لاختبارات الذكاء العام ـ سواء كانت أسئلة منفصلة أو اختبارات فرعية، وقد تناول فؤاد ابو حطب فى موضع اخر (٢٤) الأدلة والشواهد التجريبية على وجود هذه الفروق داخل الفرد الواحد.

٢ – زيادة الاعتقاد بأن ما يسمى اختبارات الذكاء العام أقل عمومية مما كان مفترضا فيها، وتوجد شواهد كثيرة على أن بعض هذه الاختبارات كمقياس ستانفورد ـ بينيه، يقيس الفهم اللفظى بالإضافة إلى عدد من القدرات الأخرى.

٣ – لم تتناول اختبارات الذكاء العام بعض جوانب النشاط العقلى مثل القدرات

الميكانيكية والموسيقية والكتابية، باستثناء قليل من الاختبارات العملية للذكاء، ولذلك سرعان ما ظهرت اختبارات «الاستعدادات الخاصة» special aptitudes ـ كما تسمى ولازالت فى بعض المؤلفات المتخصصة ـ لتكمل نقائص اختبارات الذكاء العام.

٤ – زيادة نشاط علماء النفس فى المجالات التطبيقية ـ كالتوجيه والإرشاد التربوى والمهنى، والانتقاء والتصنيف فى الميادين الصناعية والعسكرية ـ مما تطلب توافر الاختبارات التى تقيس «الاستعدادات الخاصة» مثل الاستعداد الكتابى والميكانيكى، كان انعكاسا لهذا الاهتمام.

٥ – التوسع فى استخدام نهج التحليل العاملى عند بناء الاختبارات العقلية، وقد أمكن لعلماء النفس بهذا المنهج أن يتوصلوا إلى تحديد أدق للقدرات المختلفة التى يتضمنها ذلك المصطلح العام «الذكاء». بالإضافة إلى تصنيف هذه القدرات والوصول إلى نماذج نظرية لما يسمى بالتنظيم العقلى (راجع ٢٤) ولذلك ظهرت البطاريات التى تقيس «عوامل متعددة»، ويتطلب هذا بالطبع أن تتألف من مجموعة من الاختبارات ، يقيس كل منها عاملا من العوامل التى أمكن تحديدها بالتحليل العاملى.

وسوف نتناول فى هذا الفصل طائفة من هذه البطاريات التى تقيس عددا من الاستعدادات فى وقت واحد وتسمى لهذا السبب بطاريات الاستعدادات المتعددة.

اختبارات القدرات العقلية الأولية:

تعد بطارية اختبارات القدرات العقلية الأولية من أهم نتائج استخدام منهج التحليل العاملى فى بناء الاختبارات، وقد أعد هذه البطاريات العالم الأمريكى ثرستون Thurstone محصلة لنشاطه العلمى الواسع فى ميدان تحليل القدرات العقلية، وقدنشرت فى عام ١٩٤١ باسم: اختبارات شيكاغو للقدرات العقلية الأولية The Chicago Test of Primary Mental Abilities وظلت تصدر بهذا الاسم حتى عام ١٩٤٧، وفيها انتقى ثرستون لكل عامل الاختبارات التى تشبعت به تشبعا عاليا. وأعدت لطلاب المرحلة الثانوية والجامعات وبلغ عدد الاختبارات ١١ اختبارا تقيس ٦ عوامل أولية هى: الفهم اللفظى (V) ، والقدرات المكانية (S) ، والقدرة العددية (N)، والذاكرة (M)، وطلاقة الكلمات (W)، والاستدلال (R)، واستخدم لقياس كل قدرة اختباران، ما عدا الذاكرة فقد استخدم فى قياسها اختبارا واحد، وقد طبعت هذه

التقويم النفسى ـــــــ

٣ – من الصف الرابع حتى الصف السادس الابتدائى (من ١٠ – ١٢ سنة).

٤ – تقابل المستويات العمرية للمرحلة الإعدادية فى مصر (من ١٢ – ١٥ سنة).

٥ – تقابل المستويات العمرية للمرحلة الثانوية فى مصر (من ١٥ – ١٨ سنة).

وتعطى كل بطارية من هذه البطاريات درجات منفصلة فى القدرات الأولية الآتية:

١ – الفهم اللفظى (معانى الكلمات).
٢ – القدرة العددية.
٣ – الاستدلال.
٤ – السرعة الإدراكية.
٥ – العلاقات المكانية.

وتتضمن بطارية الأعمار من ١٠ – ١٢ سنة اختبارات لقياس العوامل الخمس أما بالنسبة لبطارية المستوى الأدنى يحذف منها اختبار الاستدلال لصعوبته وبالنسبة لبطاريات المستوى الأعلى يحذف منها اختبار السرعة الإدراكية لسهولته، وتتحول درجات كل عامل، وكذلك الدرجة الكلية فى البطارية إلى نسب ذكاء انحرافية بدلا من نسبة الذكاء التقليدية فى طبعة ١٩٤٧، وتذكر كراسات التعليمات أن معلمى دور الحضانة والصفوف الأولى من المرحلة الابتدائية يهتمون بالأعمار العقلية أكثر من غيرها من المعايير ولذلك حولت جميع درجات بطارية مرحلة الحضانة حتى بداية الصف الثانى الابتدائى إلى نسب ذكاء تقليدية فقط، كما حولت درجات بطارية الصف الثانى حتى الصف الرابع الابتدائى إلى نسب ذكاء تقليدية ونسب ذكاء انحرافية وهذا المطلب واجهته أيضًا كما أشرنا مقاييس وكسلر للأطفال فى المرحلة الابتدائية WISC وفى دور الحضانة WPPSI، إلا أن هذه المقاييس أعطت الأعمار العقلية واحتفظت بانحرافات معيارية موحدة لجميع الأعمار، وهذا ما كان يجب على اختبارات القدرات العقلية الأولية عمله بدلا من تغيير وحدات القياس فى المستويات العمرية الدنيا، وبالإضافة إلى ذلك تتوافر للبطاريات الأخرى معاييرمئينية وتساعية.

وقد أثمرت بطارية القدرات العقلية الأولية ثروة من المعلومات التى توافرت من البحوث الكثيرة التى أجريت عليها فى مراحلها الثلاث، ومع ذلك فإننا نذكر

التقويم النفسى ـــــــ

القدرة العددية) وهو تمييز ليس له ما يبرره فى النظرية الحديثة للقدرات العقلية (٢٤).

وقد أعد هذه البطارية باللغة العربية أحمد زكى صالح وعدلها بما يتفق مع البيئة المصرية، وتتألف من ٤ اختبارات فرعية هى:

١ – معانى الكلمات: ويتطلب من المفحوص التعرف على الكلمة التى ترادف فى المعنى كلمة معطاة.

٢ – الإدراك المكانى: ويتطلب من المفحوص التعرف على الأشكال المنحرفة والمعكوسة أوالمقلوبة.

٣ – التفكير (أو الاستدلال): ويتطلب من المفحوص تكملة سلاسل حروف.

٤ – العدد: ويتطلب من المفحوص مراجعة صحة عمليات جمع معطاة.

وتعطى للمفحوص درجة فى كل اختبار من هذه الاختبارات الأربع، بالإضافة إلى درجة خاصة بالنتاج العام تحسب بالمعادلة الآتية:

ق.ع = ١ ل + $\frac{١}{٢}$ + ١ ف + ١ ع.

حيث الرمز ق. ع: القدرة العامة.

١ ل = درجة المفحوص فى اختبار معانى الكلمات.

١/٢ ك = نصف درجة المفحوص فى اختبار الإدراك المكانى.

١ ف = درجة المفحوص فى اختبار التفكير.

١ ع = درجة المفحوص فى اختبار العدد.

وتحول الدرجات الخمسة إلى تخطيط نفسى (بروفيل) باستخدام معايير البطارية، وقد طبقت البطارية لأكثر من ١٥ عاما على مجموعات كبيرة من الأفراد ،لايقل عدد الأفراد الذين طبق عليهم الاختبار فى كل سن عن خمسمائة فرد من الجنسين هم العينة الكلية التى استخرجت منها معايير الاختبار، وفى بعض الأعمار زاد عن ذلك مثل سن السابعة عشرة، (١ : ٩). واستخرجت المعايير لكل عمر زمنى على حدة من سن ١٢ – ١٧ سنة فى صورة مئينيات، ونسب ذكاء تقليدية.

وتبين كراسة تعليمات البطارية معاملات ثبات الاختبارات الفرعية وهى تتراوح بين ٠,٨١ ، ٠,٩٥ ، إلا أنها لم تبين حجم العينات التى حسبت منها هذه الاختبارات الفرعية، وهى بصفة عامة معاملات الارتباط بين الاختبارات الفرعية، وهى بصفة عامة معاملات منخفضة مما يدل على الاستقلال النسبى للعوامل، أما عن صدق الاختبارات فتذكر كراسة التعليمات أنه حسبت معاملات الارتباط بين الاختبارات الفرعية وغيرها من الاختبارات بالإضافة إلى التكوين العاملى للاختبارات الفرعية، ونحن ننصح المهتم بهذه البطارية بالرجوع إلى البحوث التى أشرف عليها أحمد زكى صالح حتى وفاته عام ١٩٧٤ والتى قام بها طلاب الدراسات العليا بقسم علم النفس بكلية التربية جامعة عين شمس ، للحصول على مزيد من التفاصيل، كما ننصح أى مستخدم لها فى الوقت الحاضر أن يعيد تقنين البطارية حيث أن معظم بياناتها السيكومترية تحتاج إلى تحديث.

ومن الدراسات المصرية الحديثة دراسة مرفت محمد عبد الله للحصول على درجة الماچستير فى علم النفس التربوى عام ٢٠٠٢ حيث أعادت تقنين بطارية القدرات العقلية الأولية حيث مر على تقنين هذا الاختبار فترة تجاوزت الثلاثين عاما، وعلى ذلك تتضح أهمية اعادة تقنين الاختبارات كل عشر سنوات على الأكثر.

وقد أجرى التقنين على عينة قوامها ١٥٩٨ تلميذا وتلميذة تتراوح أعمارهم بين ١٣-١٧ سنة تم اختيارهم من القاهرة الكبرى - الجيزة - القليوبية.

وقد تم تحديد معاملات صعوبة المفردات ومستوى تمييزها. ثم أجريت دراسة عاملية للأداء لتحديد مدى تشبع كل مفردة من كل اختبار فرعى (معانى الكلمات - الإدراك المكانى - الإستدلال- العدد) على العامل الخاص به.

وقد تم تحديد ثبات الاختبارات بطريقة ألفا كرونباك على ٧٨٠ مفحوصا وبلغت معاملات الثبات للإختبارات الأربعة ٠,٩٩٣١ ، ٠,٩٣٥٦ ، ٠,٩١٣٤ ، ٠,٨٨٦٧ على التوالى. أما صدق الإختبارات فقد تم تحديدها بالتحقق من صدق المفردات عامليا، وكذلك استخدم محك تمايز العمر للتحقق من أن الذكاء ينمو مع تزايد العمر.

وقد تم حساب درجات معيارية معدلة (تائية) متوسطها ٥٠ وانحرافها المعيارى ١٠ وتوافر لذلك قيم المتوسطات والإنحرافات المعيارية حتى تسهل المقارنة.

اختبارات الاستعدادات الفارقة:

ظهرت هذه البطارية لأول مرة عام ١٩٤٧ باسم اختبارات الاستعدادات الفارقة (DAT) Differential Aptitude Tests وأعدها بينيت C. K. Bennett وسيشور H. B. Seashore ووزمان A. C. Wesman ، ثم عدلت عام ١٩٦٣ وأعيد تقنينها، وظهرت الطبعة الرابعة منها عام ١٩٦٦ .

وتعد هذه البطارية من أكثر بطاريات الاستعدادات المتعددة طموحا، وتتوافر عنها - فى الخارج - بيانات إحصائية من خلال التقنين والتحليل فى دراسات عديدة كثيرة، ورغم أنها تصلح خاصة للمستويات العمرية من ١٤ سنة إلى ١٨ سنة، وفى أغراض التوجيه التربوى والمهنى، إلا أنها يمكن أن تستخدم مع الجماعات غير المنتقاة من الراشدين.

وتتألف البطارية من ٨ اختبارات قننت جميعا على نفس عينة التقنين، ومعنى ذلك أن معايير الاختبارات لها نفس الدلالة النسبية، كما أمكن التحكم فى مدى العمر الزمنى والاستعداد والمستوى الدراسى، والعوامل غير العقلية فى الشخصية، ومعنى ذلك أن التخطيط النفسى له معناه عند تفسير الفروق داخل الفرد، وتتوافر بطاريات منفصلة ومعايير مستقلة للذكور والإناث فى صورة مئينيات وتساعيات.

وتقاس القدرات المتضمنة فى هذه البطارية قياسا مستقلا، بالإضافة إلى أن الاختبارات التى تتضمنها تعتمد على القوة فى الأداء وليس السرعة، فيما عدا اختبارالسرعة والدقة الكتابية.

ورغم أن مؤلفى البطارية لم يستخدموا منهج التحليل العاملى فى إعدادها، إلا أنهم استعانوا فى الاختبارات التى تتألف منها بنتائج البحوث العاملية بالإضافة إلى المطالب العملية، ولذلك فإنهم لم يسعوا إلى اختيار الاختبارات «ذات التكوين العاملى البسيط» أو «النقية عامليا»، وإنما اختاروا الاختبارات المركبة عامليا ما دامت تفيد فى مجالات تربوية أو مهنية لها أهميتها، وتتألف البطارية من الاختبارات الآتية:

١ – الاستدلال اللفظى: وتتألف أسئلته من القياس التمثيلى اللفظى ويتشبع بالمفاهيم والعلاقات اللفظية المركبة، ومع ذلك فيغلب عليه طابع الاستدلال وليس الفهم اللفظى، لأن الاختبار لايؤكد صعوبة المفردات اللغوية، ومن أمثلة أسئلته:

... بالنسبة إلى المساء مثل الفطور بالنسبة إلى..

«أ» العشاء – الركن.

«ب» الهادئ – الصباح.

«جـ» الباب – الركن.

«د» الجرى – الاستمتاع.

«هـ» العشاء – الصباح.

ويرى فريمان (٤٩: ٤١٧) أن هذا الاختبارا يقيس فى الواقع جانبا محدودا من الاستدلال اللفظى، ولذلك يحسن أن يسمى اختبار «التماثل اللفظى».

٢ – القدرة العددية: ويشمل هذا الاختبار مدى واسعا من العمليات الحسابية، ولا يلعب فيه حل المسائل الحسابية أى دور، وتقيس بعض الأسئلة المهارة فى إجراء العمليات الأساسية الأربعة، ويتطلب البعض الآخر فهم المفاهيم والعلاقات الكمية.

٣ – الاستدلال المجرد: ويقيس القدرة على الاستدلال باستخدام المواد غير اللفظية، ويتألف من سلاسل من الرسوم، يتطلب كل منها فهم المبدأ الذى يؤدى إلى التغير فى الأشكال التى تتألف منها السلسلة أى من نوع سلاسل الأشكال.

٤ – العلاقات المكانية: ويقيس القدرة على التصور البصرى المكانى وذلك بعرض شكلين هندسيين من الأشكال الثنائية البعد، ويقوم المفحوص بمعالجتهما ذهنيا بحيث يؤديان إلى شكل ثلاثى البعد، والغرض من الاختبار قياس القدرة على التصور البصرى للأشكال المركبة بعد تدويرها.

٥ – الاستدلال الميكانيكى: ويقيس القدرة على الفهم الميكانيكى بواسطة مواقف مصورة تمثل المبادئ الميكانيكية والعملية، ويصاحب كل صورة سؤال بسيط موجز حول المبدأ المتضمن.

٦ – السرعة والدقة الكتابية: ويقيس السرعة والدقة فى الاستجابة للاقترانات الثنائية بين الحروف والأعداد، ويتطلب المزاوجة بين مختلف الاقترانات مؤكدا إدراك التفاصيل ومعدل الاستجابة، وتهدف الأسئلة التى يتكون منها إلى «الاقتراب من

العناصر التى تتألف منها كثير من المهن الكتابية».

٧ – الاستخدام اللغوى (القسم الأول): ويقيس القدرة على تهجى الكلمات وذلك عن طريق تحديد ما إذا كانت الكلمات التى يتألف منها الاختبار صحيحة أم خاطئة من الناحية الإملائية.

٨ – الاستخدام اللغوى (القسم الثانى): ويقيس القدرة على استخدام القواعد اللغوية، وذلك عن طريق تحديد ما إذا كانت الجمل التى يتألف منها الاختبار صحيحة أم خاطئة من ناحية النحو وعلامات التوقيف واستخدام الكلمات.

وتتكون هذه البطارية من صورتين متكافئتين هى الصورة (ل) والصورة (م)، وتعطى كراسة التعليمات الخاصة بالطبعة الرابعة التى ظهرت عام ١٩٦٦ تفاصيل كافية عن الإجراءات التى استخدمت فى بناء البطارية، وقد قننت فى الولايات المتحدة على ٥٠٠٠٠ تلميذا فى ١٩٥ مدرسة فى ٤٣ ولاية.

ويذكر السيد محمد خيرى (١٥ : ٣٣١) أنه قام بترجمة هذه البطارية إلى اللغة العربية بالاشتراك مع لويس كامل مليكة. وتستخدم فى مصر أجزاء مختلفة كثيرة من هذه البطارية لأغراض متعددة منها اختبار الاستدلال الميكانيكى الذى قامت مراقبة الاختبارات السيكولوجية بمصلحة الكفاية الإنتاجية باقتباس جزء منه وتعديله بما يتناسب مع أهداف الانتقاء لنظام التلمذة الصناعية فى حرف مختلفة، واختبار الاستدلال اللفظى والسرعة والدقة فى الأعمال الكتابية اللذان اقتبسا وعدلا عند استخدامها فى اختبارات ديوان الموظفين الذى كان موجودا فى مصر خلال الخمسينات والستينات من القرن الماضى.

والواقع أنه لاتتوافر لنا أى بيانات تقنين عن هذه البطارية ككل فى اللغة العربية فيما عدا اختبار الاستدلال الميكانيكى، فقد حسبت المئينيات والدرجات التائية لعينة حجمها ٢٣٣٩ تلميذا فى مراكز التدريب المهنى التابعة لمصلحة الكفاية الإنتاجية والتدريب المهنى لوزارة الصناعة (١٦ : ٤٩).

كما حسب له معامل ثبات بطريقة التجزئة النصفية (١٦ : ١٥) فبلغ ٠,٦١ واستخدم فى حساب صدقه ـ كجزء من بطارية اختبارات استعدادات حرف المعادن التى سنشير إليها فيما بعد ـ منهج التحليل العاملى فبلغ مقدار تشبعه «العامل الميكانيكى

العام، ٠,٥٣ ، ولم يحسب معامل ارتباطه بمحك من محكات التدريب المهنى (١٦ : ٢١).

بطارية الاستعدادات العامة:

ظهرت بطارية الاستعدادات العامة General Aptitude Test Battery (GATB) عام ١٩٤٧ وأعدها خبراء مكتب التوظيف الأمريكى لأغراض التوجيه والإرشاد فى ميدان الخدمة المدنية، وتعتمد على افتراض أن عددا كبيرا من الاختبارات يمكن أن يتجمع فى عدد من العوامل وأن عددا كبيرا متنوعا من المهن يمكن أن يتجمع فى فئات مهنية تبعا لدرجة التشابه بينها فى القدرات المطلوبة، وييسر هذا اختبار جميع القدرات المهنية عند الشخص فى جلسة واحدة وتفسير درجاته فى ضوء مدى واسع من المهن، (٣٧: ٤٢٣).

وتتألف البطارية من ١٢ اختبارات تعطى ٩ درجات عاملية، وتتطلب حوالى ساعتين ونصف فى الإجراء، والعوامل التسع التى تحددت بمنهج التحليل العاملى:

١ – الذكاء العام: ويقاس بالدرجة المركبة من ٣ اختبارات هى: الإدراك المكانى الثلاثى، والمفردات اللغوية، والاستدلال الحسابى.

٢ – الاستعداد اللفظى: ويقاس باختبار المفردات الذى يتطلب من المفحوص التعرف على معانى الكلمات (مترادفات) وأضدادها.

٣ – الاستعداد الحسابى: ويشمل أسئلة تتطلب إجراء العمليات الحسابية وحل المسائل الحسابية من النوع المعتاد.

٤ – الإدراك المكانى: ويقيس القدرة على التصور المكانى للأشياء ذات الأبعاد الثلاثة حين تعرض عرضا ثنائيا.

٥ – إدراك الشكل: ويقيس القدرة على المزاوجة بين رسوم لآلات معينة من ناحية وأشكال هندسية من ناحية أخرى.

٦ – الإدراك الكتابى: ويتطلب المزاوجة بين الأسماء،

٧ – التآزر الحركى: ويقيس القدرة على التآزر، وذلك بأن يطلب من المفحوص وضع علامات معينة بالقلم الرصاص فى سلسلة من المربعات.

٨ – مهارة الأصبع: ويقيس كفاءة المفحوص فى تجميع الأشياء الصغيرة وفكها.

٩ – المهارة اليدوية: ويقيس المهارة اليدوية باستخدام اليدين معا، ثم اليد المفضلة، فى وضع العصى الخشبية فى لوحة العصى، ثم قلب أوضاع العصى.

وتتطلب الاختبارات الأربع التى تقيس عاملى مهارة الاصبع والمهارة اليدوية استخدام بعض الأجهزة البسيطة، أما الاختبارات الثمانى الأخرى فهى من نوع الورقة والقلم. وتوجد صور متكافئة للاختبارات السبع الأولى التى تستخدم فى قياس العوامل الست من الذكاء العام حتى الإدراك المكانى.

وتتحول الدرجات العاملية التسع فى بطارية الاستعدادات العامة إلى درجات معيارية معدلة متوسطها ١٠٠ وانحرافها المعيارى ٢٠، واشتقت هذه المعايير الأمريكية من عينة حجمها ٤٠٠٠ مفحوصاً تمثل المجتمع الإحصائى السكانى العام من حيث العمر والجنس والمستوى التعليمى والتوزيع المهنى والجغرافى، وباختبار جماعات من الموظفين والمتقدمين للوظائف والمتدربين فى برامج التدريب لمختلف المهن، وقد أمكن تحديد أنماط الدرجات التى تبين الاستعدادات الفارقة، والحد الأدنى المطلوب من الدرجة المعيارية لكل مهنة من المهن، فمثلا وجد أن مهنة المحاسبة تتطلب درجة ذكاء مقدارها ١٠٥، ودرجة استعداد عددى مقدارها ١١٥، أما مهنة السمكرة فتتطلب أن يكون الحد الأدنى من الذكاء ٨٥ ودرجة مقدرها ٨٠ فى الاستعداد العددى والاستعداد المكانى والمهارة اليدوية، وهكذا يمكن المقارنة بين بروفيل الدرجات المعيارية للمفحوص فى جميع المهن التى يصل إلى درجات القطع عندها أو يتفوق عليها ويستفاد بهذه المقارنة فى أغراض التوجيه والإرشاد المهنى.

وقد قامت وحدة بحوث الدراسات النفسية بالمركز القومى للبحوث الاجتماعية والجنائية (٣٥) بنقل هذه البطارية إلى اللغة العربية، وتم اختيار ٧ من اختبارات البطارية الاثنى عشر لتقنينها على عينات مصرية هى: اختبار العدد (العمليات الحسابية)، واختبار إدراك العلاقات المكانية ثلاثية البعد، والمفردات، واختبار المزاوجة بين الأدوات، واختبار الاستدلال الحسابى. واختبار المزاوجة بين الأشكال، واختبار رسم العلامات،وطبقت على عينة من تلاميذ السنة الثالثة الإعدادية، وتلاميذ المدارس الثانوية العامة والثانوية الفنية، وتلاميذ مراكز التدريب المهنى التابعة لوزارة الصناعة

بلغ عددها ٤٩٨ تلميذا، واستخدمت درجات هؤلاء التلاميذ فى تحليل الأسئلة التى تتألف منها الاختبارات السبع فحسبت معاملات التمييز، ودرجات الصعوبة، ثم طبقت البطارية فى صورتها المعدلة على عينة تقنين كبيرة نسبيا بالمقارنة ببحوث التقنين التى أجريت على الاختبارات المصرية من الصف الثالث الإعدادى والصف الثالث الثانوى (علمى وأدبى)، والتعليم التجارى والصناعى، والزراعى، ومراكز التدريب المهنى التابعة لوزارة الصناعة.

وحسبت معاملات الثبات بطريقة إعادة الاختبار فتراوحت بين ٠,٧٩ ، ٠,٩٣ بالنسبة لمختلف الاختبارات الفرعية داخل نطاق أنواع التعليم المختلفة، كما استخدمت طريقة التجزئة النصفية لاختبارات القوة فتراوحت معاملات الثبات بين ٠,٨١ ، ٠,٩٣ ، وطريقة كيودر - ريتشاردسون لاختبارات السرعة فتراوحت معاملات ثباتها بين ٠,٧٥ ، ٠,٩٦ وحسبت المئينيات للعينات الفرعية التى تتألف منها عينة التقنين الكلية.

وقد حسب صدق التكوين الفرضى عن طريق التحقق من قدرة اختبارات البطارية على التمييز بين أداء تلاميذ الفئات المختلفة من التعليم، واستخدمت لهذا الغرض مصفوفة متسلسلة من النسب الحرجة لتحديد قيمة الدلالة الإحصائية للفروق بين المتوسطات - كما حسب الصدق التنبؤى على عينة من ٢٠٠ تلميذ من مراكزالتدريب المهنى بأنواعها المختلفة وقورنت درجاتهم فى البرامج التدريبية التى تعرضوا لها قبل تطبيق الاختبارات عليهم وبعده، وعرضت النتائج فى صورة جداول توقع لكل اختبار على حدة، ومعنى ذلك أن درجات هذا الاختبار يمكن أن تفسر فى ضوء كل من المعيار والمحك.

اختبارات فلاناجان لتصنيف الاستعدادات:

تمثل هذه البطارية اتجاها مختلفا فى بناء بطاريات الاستعدادات المتعددة، وتسمى Flanagan Aptitude Classification Tests (FACT).

وقد ظهرت نتيجة لبحوث فلاناجان حول إعداد اختبارات تصنيف أفراد القوات الجوية بالجيش الأمريكى أثناء الحرب العالمية الثانية، ولذلك فإنها تتجه وجهة عملية تطبيقية، وتهتم بمجال التوجيه المهنى: وقد اعتمدت على منهج تحليل العمل لكثير من

المهن الذى أدى إلى تحديد ٢١ «عنصرا مهنيا حاسمًا» أو قدرة تميز بين الناجحين والفاشلين من العاملين فى كل مهنة. ويتميز أى عنصر مهنى أو قدرة بأنه على قدر من العمومية، أى يشترك فى أنماط عديدة من المهن، وفى الوقت نفسه على قدر من النوعية بمعنى أنه يقيس شيئا مختلفا من العناصر المهنية الأخرى.

وتتألف البطارية من ٢١ اختبارا لقياس العناصر المهنية منها ١٩ اختبارا كتابيا،واختباران عمليان هما النقش Carving والنقر Tapping، وهذه الاختبارات تمثل بطارية غير متجانسة تتضمن وظائف سلوكية مختلفة، منها الإدراك الحسى والنشاط الحركى والذاكرة والنشاط اللغوى والعد والمفردات اللغوية العامة والاستدلال اللغوى والاستخدام اللغوى، وفيما يلى عرض موجز لهذه العناصر التسعة عشر التى تقيسها الاختبارات اللفظية.

١ – الفحص inspection: وتقيس القدرة على تحديد الأخطاء بسرعة ودقة فى سلسلة من رسوم الأشياء (وهو اختبار للإدراك البصرى للتفاصيل).

٢ – التجميع Assembly : ويقيس القدرة على التصور البصرى لشىء معين حين توضع مجموعة من الأجزاء معا (وهو اختبار لإدراك العلاقات المكانية الثلاثية البعد).

٣ – الحكم والفهم Judgment & comprehension: ويقيس القدرة على القراءة مع الفهم واستخدام الحكم السليم فى المواقف العملية (اختبار لفهم معانى الفقرات).

٤ – الحذق ingenuity: ويقيس القدرة على الابتكار والاختراع فى إعداد إجراءات أو آلات أو عروض ذكية (اختبار استدلالى يعتمد على مشكلة محددة).

٥ – اليقظة alertness : ويقيس القدرة على تقدير الموقف وملاحظة ما ينشأ فيه من أخطار أو ما يحتاجه من وسائل وخطط (اختبار لإدراك التفاصيل فى صورة معينة وما بينها من علاقات).

٦ – الترميز coding: ويتطلب السرعة والدقة فى ترجمة معلومات مكتبية إلى لغة شفوية معينة.

٧ – الذاكرة memory: ويقيس القدرة على استدعاء الرموز المتعلمة من

اختبارالترميز السابق.

٨ – الدقة Precision: ويتطلب السرعة والدقة فى إصدار حركات دقيقة للأصابع.

٩ – المقاييس Scales: ويتطلب السرعة والدقة فى قراءة المقاييس والرسوم البيانية.

١٠ – التآزر Coordination : ويقيس القدرة على التآزر بين حركات اليد والذراع.

١١ – الحساب arithmetic : ويتطلب الكفاءة فى العمليات الحسابية الأساسية.

١٢ – الأنماط Patterns : ويقيس القدرة على محاكاة واسترجاع الخطوط الرئيسية لنمط بسيط مرسوم.

١٣ – المكونات Components: ويقيس القدرة على تحديد الأجزاء المهمة فى مجموعة من الرسوم والتخطيطات،

١٤ – الجداول tables: ويقيس القدرة على قراءة نوعين من الجداول أولهما يستخدم الأعداد، والآخر يستخدم الكلمات والحروف الأبجدية.

١٥ – الميكانيكا mechanics: ويقيس القدرة على فهم المبادئ الميكانيكية.

١٦ – التعبير expression: ويقيس القدرة على نقل الأفكار كتابة أو شفويا.

١٧ – الاستدلال reasoning: ويقيس القدرة على حل المشكلات والتعبير عن الحل فى صورة رياضية.

١٨ – المفردات vocabulary : ويقيس القدرة على فهم معانى الكلمات.

١٩ – التخطيط planning: ويتطلب الكفاءة فى تنظيم البيانات وذلك بإعادة ترتيب خطوات معينة لحل المشكلة.

ورغم أن الاختبارات التى تتألف منها البطارية (وعددها ٢١ كما ذكرنا) تصحح منفصلة . فإن درجات مختلف الاختبارات تجمع بعضها إلى بعض عند التنبؤ بالنجاح فى مهن معينة، وتتفاوت هذه المهن من المستوى المهنى الرفيع (كالطب والهندسة

والتدريس) وإلى مستوى العامل الماهر. وحيث أن هذه الاختبارات المنفصلة تمثل تنوعا واختلافا فى العمليات النفسية التى تقيسها، بالإضافة إلى انخفاض معاملات الارتباط فيما بينها، فإن جمع درجات هذه الاختبارات يصبح ممارسة تثيرالكثير من التساؤل والشك.

وتتطلب بطاريات الاختبارات اللفظية زمنا كليا طوله عشر ساعات ونصف تقسم إلى ٣ جلسات اختبارية، وقد أعدت معاييرها الأمريكية على عينة كبيرة من طلاب المرحلة الثانوية، ومع ذلك فإنها يعوزها التمثيل كما تقول انستازى (٤٠: ٣٤٢).

ويرى فلاناجان أن منهج العناصر المهنية يقع فى منزلة متوسطة بين تحليل القدرات إلى عوامل أولية (أو نقية) من ناحية وطريقة عينة العمل التى تتم فيها المماثلة بالعناصر الأساسية فى الأعمال والمهن الحقيقية من ناحية أخرى.

وتؤكد نتائج البحوث التى قام بها فلاناجان (٤٧) أن معاملات الارتباط بين الاختبارات الفرعية منخفضة بوجه عام حيث يبلغ وسيطها ٠,٢٠ فى عينة من الطلاب من سن ١٥ سنة، و٠,٣١ فى عينة من الطلاب من سن ١٨ سنة، ومعنى هذا أن الصدق الداخلى للبطارية قد تحقق (أى الاستقلال النسبى للاختبارات الفرعية).

وتؤكد نتائج الصدق التلازمى باستخدام محك التحصيل المدرسى عدم اللجوء إلى استخدام الاختبارات الفرعية بسبب انخفاض معاملات الارتباط بصفة عامة، كل على حدة، وفى أغراض التوجيه التعليمى، والواقع أن الدرجة المجمعة لدرجات جميع الاختبارات أدت إلى معاملات ارتباط متعددة لها دلالة إحصائية، إلا أن الصعوبة الجوهرية فى اللجوء إلى البطارية ككل تكمن فى الوقت الطويل الذى تحتاجه فى تطبيقها.

وقد قام فلاناجان بدراسات تتبعية لأفراد عينة من طلاب الجامعة فى عدد من مجالات الدراسة، بعد أن اختبرهم فى أثناء دراستهم الثانوية (لحساب الصدق التنبؤى)، وبلغ وسيط معاملات الارتباط ٠,٣٩ وكانت أقل معاملات الارتباط لطلاب الدراسات الدينية والخدمة الاجتماعية وأعلاها لطلاب العلوم الاجتماعية (التى شملت ٦ منها على سبيل المثال الأخصائى النفسى ورجل القانون، والمؤرخ)، كما حسبت معاملات الارتباط بين درجات الاختبارات الفرعية وتقديرات الرؤساء والمشرفين فى

عدد من المهن ، فوجد أن هذه المعاملات منخفضة بصفة عامة، بل قد تصل فى بعض الأحيان إلى الصفر أو عدم الدلالة الإحصائية.

وقد أجريت على هذه البطارية دراستان فى مصر أحدهما على طلاب كلية الهندسة (٢ أ) والأخرى على طلاب كلية الطب (٣٥ أ).

الإختبارات العملية

من الأخطاء الشائعة فى مجال القياس النفسى افتراض وجود ارتباط دال بين معرفة الفرد للمعلومات والحقائق والمباديء والقوانين المرتبطة بسمة معينة من جهة والأداء العملى الفعلى المرتبط بذات السمة من جهة أخرى، وعليه يتم الإهتمام باحد هذه الجوانب (النظرية – العملية) عند قياس السمة . فعلى سبيل المثال اختبارات القبول التى تستخدم عند انتقاء الطلاب فى مجال التدريس أو المجال الطبى أو مجال اعداد الطعام أو الخياطة حيث يتم الإعتماد على اختبارات لفظية دون الاهتمام بالجوانب العملية.

وبالطبع إن الاختبارات اللفظية التى نقيس من خلالها معلومات الطالب لها أهمية كبيرة سواء فى مجال التعليم أو الصناعة، وبالرغم من أهميتها إلا أنها غير كافية حيث أنها تظهر جانبا واحدا فقط فى المجال، ولكن هناك مواقف أكثر تعقيدا تتطلب من المتعلم أن يقوم بها مثل العزف على آلة موسيقية، أو تركيب أداة معينة أو اعداد وجبة طعام معينة وهذه الأعمال لا تتطلب وصفا لفظى فقط وإنما أداء فعلى للوصول إلى المنتج النهائى.

وبالمثل إن معرفة اسم ومكان الفرامل أو دواسة السرعة أو التروس المحركة أو ناقل الحركة بالإضافة الى الترتيب الزمنى لتشغيل هذه الاجزاء. إلا أن هذا لا يؤكد أن الفرد يستطيع أن يقود السيارة. وبالمثل أيضا إن معرفة الفرد لعدد كبير من الألفاظ والقواعد اللغوية لا يعنى أن الفرد يستطيع أن يعبر عن نفسه أو أن يكتب قصة قصيرة.. وهكذا.

وعلى ذلك تتضح أهمية وجود الإختبارات اللفظية إلى جانب الاختبارات العملية عندما يتطلب الأمر ذلك عند قياس إحدى السمات، ومن أمثلة الاختبارات العملية المألوفة فى قياس القدرة العقلية وضع بعض الأشكال الهندسية فى أماكنها المناسبة فى

لوحة ما، أو اختبارات المتاهات. وتستخدم هذه الإختبارات العملية للحكم على القدرة العقلية للأفراد الذين يعانون من ضعف فى اللغة، أو للأفراد الذين لم ينالوا قسطا من القراءة والكتابة، ومن هذا المنطلق نجد أن الإختبارات العملية تكون مرادفة للإختبارات غير اللفظية، وهذه الإختبارات لا تستخدم كبديل لإختبارات الذكاء ولكن اضافة إلى اختبارات الورقة والقلم الخاصة بالقدرات اللفظية.

ويمكن من خلال الاختبارات العملية قياس المهارات العملية فى مجال الرياضة البدنية كالتصويب نحو هدف معين، أو العزف على الالات الموسيقية والتى تتطلب سرعة ودقة فى أدائها أو عند قياس التآزر الحركى والعضلى والعصبى مثل المهارة الإصبعية Finger dixterity ، والمهارة اليدوية Manual Dexterity وهى مهارات لازمة بالتنبؤ بالقدرة على اكتساب المهارات الحركية للنجاح فى مجالات تربوية معينة مثل التعليم الصناعى أو الإقتصاد المنزلى أو الكتابة على الآلة الكاتبة أو عزف آلة البيانو – على سبيل المثال لا الحصر – وهذه المهارات تتطلب تدريبا متواصلا حتى يحدث التقدم فى السلوك المهارى المطلوب احرازه. ويتطلب ذلك بالطبع تحديدا دقيقا بطبيعة المهمة التى نسمى اختيار الأدوات الملائمة لقياسها، وإلى أى مدى يمكن التنبؤ بالسلوك فى المواقف اللاحقة فى ضوء السلوك المقاس، وأيضا إلى أى مدى يمكن الإعتماد على هذه الأدوات فى انتقاء الأفراد وتصنيفهم، ولا يمكن الوصول إلى ذلك إلا من خلال التحليل العلمى للعمل والذى يمكننا من تحديد الصدق التآليفى أو التركيبى Synthetic validation (انستازى وأوريانا ٤٠ : ٤٩٥) وتعرف اناستازى وأوريانا مصطلح الصدق التأليفى بأنه «الصدق الذى يستدل عليه فى موقف معين من خلال التحليل الموضوعى لمكونات العمل، والتحقق من صدق هذه المكونات، وتجميع هذه المكونات فى كل». وأرجعتا هذا التعريف الى بالما Balma (٤٣ : ٣٩٥) ويتضمن ذلك القيام بخطوات ثلاث هى:

١– تحليل مفصل للعمل أو المهمة للتعرف على مكونات العمل وأوزانها.

٢– تحليل ودراسة ميدانية لكل أداة أو اختبار لتحديد إلى أى مدى يقيس المكونات المتضمنة فى الاداء وكفاءته.

٣– تحديد صدق كل أداة أو اختبار استخدم فى قياس مكونات العمل وأوزانها.

وقد حدد بريموف Primoff: ١٩٥٩ ، بريموف وهايد (١٩٨٨:٥٨) طريقة تحديد معامل الصدق التأليفى أو التركيبى المرتبط بتحليل العمل Job validity ويرمز له J. validity بتعديل لطريقة حساب الإنحدار المتعدد.

أنواع الإختبارات العملية:

هناك أنواع متعددة من الاختبارات العملية والتى تستخدم بصفة عامة فى المجالات المختلفة من التعليم أو الصناعة والتى تستخدم فى أغراض متعددة مثل الإنتقاء أو التصنيف أو تقويم الأداء النهائى للعمل وكفاءته أو كفاءة العاملين أو المتقدمين لوظائف تتطلب مهارات معينة. ويمكن تقسيم الإختبارات العملية إلى فئات ثلاث هى:

١- اختبارات التعرف Recognition Tests

٢- اختبارات المواقف المتشابهة Simulated condition

٣- اختبارات عينة العمل Work Sample Test

وفيما يلى وصفا مختصرا لكل نوع

١- اختبارات التعرف:

وتهدف إلى قياس قدرة الفرد على التعرف على خصائص المنتج النهائى للعمل أو الخصائص اللازمة للعمليات الواجب القيام بها لإنتاج العمل، أو تحديد الاخطاء التى وقع فيها شخص ما عند انتاج العمل، أو قد يطلب من المفحوص أن يبين كل جزء من اجزاء العمل أو آلة مستخدمة فى انتاج العمل مع قدرة المفحوص على تحديد وشرح العمل .

٢- الإختبارات التى تتضمن مواقف مشابهة:

فى بعض الأحيان تجرى الإختبارات على مواقف أو عمليات متشابهة للمواقف الأصلية حيث يكون الإختبار وسيلة القياس، وهذه الاختبارات يجب أن تتضمن النشاطات الضرورية للقيام بالعمل، ويلاحظ المختبر أو المشرف على العمل ما يقوم به المفحوص من أعمال متشابهة للموقف الأصلى أو موقف مصغر منه مثل التدريس المصغر للطالب المعلم.

وبالرغم من أن استخدام المواقف المشابهة لها بعض الفوائد إلا أن هذه الطريقة لابد أن تؤخذ بحذر وبمعرفة كافية للعلاقة بين نتائج الإختبار المصغر والمعالجة الكاملة تحت الظروف الحقيقية.

٣- اختبارات عينة العمل:

فى هذا النوع من الاختبارات يجب أن يشتمل الاختبار على عينة حقيقية للعمل تحت الظروف والشروط الحقيقية لإنجاز العمل. وهذا النوع من الإختبارات لها درجة عالية من الصدق عن النوعين الآخرين.

ويوجد نوعان أساسيان فى هذه الفئة من الإختبارات هى:

أ- اختبارات يمكننا من خلالها تمييز الصواب أو الخطأ عند تنفيذ المهارة ويمكن فى هذه الحالة تحديد الدرجات أوتوماتيكيا مثل اختبارات التصويب نحو هدف معين .

ب- اختبارات تتطلب أحكام الملاحظين أثناء العمل وتحديد الدرجة ومستوى الاداء مثل قيادة السيارة أو عزف آلة موسيقية.

استخدامات اختبارات الأداء

توجد أهداف مباشرة لإستخدام الاختبارات العملية، ألا وهى قياس الفروق بين الأفراد ومدى التقدم فى عمل ما، أما الهدف النهائى فهو العائد النهائى من التعليم والتدريب. وتنقسم هذه الإستخدامات الى:

١- التنبؤ بإجراء المهارات بنجاح.

٢- تشخيص نقاط القوة والضعف فى الأداء.

٣- المساعدة فى التدريس.

٤- ايجاد محكات للقياس.

ففي الاستخدام الأول يتم تحديد مدى تقدم الأفراد بالنسبة لطريقة أو عملية معينة، ويعتبر هذا التقدير من أهم الإستخدامات فى التعلم والصناعة والحياة اليومية.

أما الإستخدام الثانى والذى يهدف الى تشخيص الصعوبات فى الأداء حيث يقدم الإختبار العملى شروطا مقننة لكل أداء أو النقص فيه. ومن خلال ذلك يمكن تصويب

الأخطاء وتعديل السلوك للوصول إلى الهدف المرتبط بالعمل.

واستخدام الاختبارات العملية فى العملية التعليمية يرتبط بعاملين أولهما معرفة النتائج المرتبطة بالأداء حيث تستثير دوافع المتعلمين من خلال معرفة المتعلم بنتائج أداءه من حيث النجاح أو الإخفاق فى تنفيذ العمل.

أما بالنسبة للمعلم فإن نتائج أداء المتعلم لما هو مطلوب منه تؤدى إلى تعديل وتقويم طريقة التدريس والتى تمكن المتعلمين من اكتساب مهاراتهم فى الإتجاه المرغوب فيه.

ويعتبر تطبيق الإختبارات العملية ونتيجة الاداء فيه هو المرجع لوضع المتعلم فى المستوى الصحيح سواء فى أثناء التدريب أو عند ممارسة العمل الفعلى، حيث تكون النتائج بمثابة المحك للتقويم العام فى البرامج التدريبية فى مجال التعليم أو الصناعة حينما يكون الهدف النهائى الاساسى للنشاط التعليمى هو العمل على تنمية الاداء المهارى.

تقويم الاداء فى الإختبارات العملية:

يعتمد تقويم الأداء فى الإختبارات العملية على عناصر العمل المطلوب تعلمه أو التدريب عليه، وكذلك على الوزن النسبى للعناصر المطلوب أدائها. ويتم تقدير الدرجات بناء على مقياس تقدير يحدد الدرجة الفعلية التى أحرزها المتعلم فى كل عنصر من عناصر العمل مراعين فى ذلك سرعة الأداء ودقته والزمن المطلوب لإنجاز كل وحدة فى العمل.

ثبات الإختبارات العملية:

إن الطريقة المثلى لتحديد ثبات الإختبارات العملية هو ثبات المصححين حيث يقوم كل فاحص بتقدير أداء المفحوص مستقلا عن الفاحص الآخر، ويتم حساب معامل الثبات والذى يطلق عليه معامل الإتفاق بين المصححين. وقد يتم حساب الثبات من خلال فحص العمل النهائى، وكلا الطريقتين مقبولتين وإن كانتا متكاملتين، ففحص الأداء النهائى لا يوضح طريقة العمل التى أدت الى عدم كفاءة المنتج أو الوقوع فى خطأ معين. كذلك يتطلب الأمر وصف الأداء النموذجى وتحديد

نقاط قطع للدرجات بحيث يحدد بدقة متى يحصل المفحوص على درجة معينة.

وعلى ذلك تتحدد الخطوات التى يتم فيها ملاحظة الاداء فى الموقف الاختبارى كالآتى:

١- اختيار الملاحظين المدربين الذين سيقومون بملاحظة الموقف الإختبارى.

٢- التدريب على ما سيتم ملاحظته وأيضا كيفية تقدير الدرجات.

٣- التعرف على مفردات المقياس (مقياس التقدير المستخدم والذى يتضمن العناصر التى سيتم تقديرها) وذلك للتأكد من وجود اتفاق بين الملاحظين على المفردات. وفى حالة عدم وجود بطاقات ملاحظة ملائمة للموقف الملاحظ فيجب تصميم بطاقة ملاحظة لتقدير مختلف عناصر الأداء والتى يجب تحديدها بدقة وموضوعية مع وضع التعليمات الخاصة باستخدام هذه البطاقة وبصفة خاصة عندما يتطلب الأمر ملاحظة وتقدير الأداء اثناء التدريب أو اثناء أداء العمل النهائى والذى يصعب فيهما اعادة الموقف لتقديره.

بطارية اختبارات استعدادات حرف المعادن:

أعدت بطارية اختبارات استعدادات حرف المعادن فى مصلحة الكفاية الإنتاجية والتدريب المهنى بوزارة الصناعة بإشراف السيد محمد خيرى (١٣)، وتطلب إعدادها تحليل الأعمال التى يدرب عليها تلاميذ مراكز التدريب المهنى التابعة لهذه المصلحة، وأسفرت نتائج هذا التحليل بالنسبة لجميع حرف المعادن والسيارات والتبريد على استخلاص القدرات المهنية التالية: العلاقات المكانية – الفهم – الميكانيكى – التصور البصرى، تذكر الأشكال، سرعة الإدراك، المعلومات أو الخبرة الميكانيكية، تقدير الأطوال والمساحات والحجوم، ثبات اليد، مهارة الأصابع، وأعدت بطارية اختبارات لقياس هذه القدرات تتألف من قسمين أحدهما يعتمد على ورقة وقلم والآخر من النوع العملى، وفيما يلى وصف موجز لهذه الاختبارات.

اختبارات الورقة والقلم: وتشمل ٨ اختبارات هى:

١ – الاستدلال اللفظى: ويقيس القدرة على الاستدلال اللفظى وإدراك المفاهيم اللفظية والقدرة على التجريدويتألف من ٣٠ سؤالا ويستغرق ١٥ دقيقة، ويتكون كل

سؤال من جملة تتضمن مفهوما أو علاقة معينة على المفحوص اكتشافها بالاختيار من متعدد.

٢ – اختبار الذكاء الإعدادى: للسيد محمد خيرى.

٣ – اختيار الاستدلال الميكانيكى: وهو مقتبس من بطارية اختبارات الاستعدادات الفارقة التى أشرنا إليها آنفا.

٤ – اختبار المعلومات الميكانيكية، ويهدف إلى معرفة مدى ألفة المفحوص بالعدد والآلات الميكانيكية ويتكون من ٥ أسئلة فى كل منها ست صور لعدد ميكانيكية بسيطة، ويطلب من المفحوص التعرف على أسمائها بالاختيار من متعدد.

٥ – اختبار المعلومات الحسابية: ويقيس القدرة على التعامل مع الأعداد وذلك بإجراء العمليات الحسابية الأساسية (الجمع والطرح والضرب والقسمة باستخدام الأعداد الصحيحة والكسور الاعتيادية).

٦ – اختبار التصور المكانى: وهو مقتبس من بطارية الاستعدادات العامة التى أشرنا إليها آنفا.

٧ – اختبار تكميل الأشكال: وهو مقتبس من بطارية المعهد القومى البريطانى لعلم النفس الصناعى ويقيس القدرة على إدراك العلاقات الميكانيكية الثنائية البعد، والقدرة على المزاوجة السريعة بين الأشكال وإصدار الأحكام عليها بطريقة تعتمد على التفكير الاستنباطى وسرعة الإدراك وتقدير المساحات، ويتطلب فى الأسئلة الصعبة التصور البصرى.

٨ – اختبار تذكر الأشكال: من إعداد السيد محمد خيرى، ويقيس القدرة على تذكر الأشكال الهندسية بسرعة.

الاختبارات العملية: وتشمل ٤ اختبارات هى:

٩ – اختبار التجميع الميكانيكى: وهو مقتبس من أصل ألمانى، ويقيس القدرة الميكانيكية وخاصة القدرة على إدراك العلاقات والمتعلقات الميكانيكية والتصور المكانى، ويعتمد على اختبارات مينيسوتا للقدرة الميكانيكية، ويعتمد على تجميع الأشياء.

١٠ – اختبار مهارة الأصابع: وهو من تصميم محلى يجمع بين جهاز أو أكثر

لمهارة الأصابع اختيرت من عدد من الاختبارات الموجودة بالفعل، ويقيس القدرة على تحريك الأصابع وتناول الأشياء الدقيقة بسرعة ودقة.

١١ – اختبار ثبات اليد: ويقيس القدرة على التحكم فى حركة اليد.

وقد قننت هذه البطارية على عينة كبيرة تتراوح أعمارها بين ١٤، ١٨ سنة من الحاصلين على الشهادة الإعدادية، وكان عددها «ينحصر بين ألف وثلاثة آلاف تلميذ على حسب الظروف التى طبقت فيها الاختبارات المختلفة» (٢ : ١٠) «وشملت تلاميذ المدارس الإعدادية والثانوية والمتقدمين لمراكز التدريب المهنى وتلاميذ هذه المراكز، وحللت أسئلة الاختبارات وتحددت لها معاملات التمييز، ورتبت حسب مستويات الصعوبة.

وقد حسب الثبات لمعظم الاختبارات بطريقة التجزئة النصفية، أما بالنسبة للاختبارات التى تعتمد على السرعة وهما اختبارا مهارة الأصابع وثبات اليد فقد استخدمت معها طريقة إعادة الاختبار، مع ملاحظة أن اختبار العمليات الحسابية لم يبين معامل ثباته فى التقرير المنشور، وكانت معاملات الثبات مرضية بصفة عامة.

ولحساب صدق التكوين الفرضى استخدم منهج التحليل العاملى لاختبارات الورقة والقلم فقط واستخرج عامل عام يكشف عن مدى تشبع كل اختبار بالاستعداد الميكانيكى العام وكانت التشبعات على النحو الآتى:

الاستدلال اللفظى ٠,٥٧ والذكاء الإعدادى ٠,٥٨، الاستدلال الميكانيكى ٠,٥٣ المعلومات الميكانيكية ٠,٣٧ والعمليات الحسابية ٠,٣٧ التصور المكانى ٠,٥٨ تكميل الأشكال ٠,٥١ تذكر الأشكال ٠,٤٧.

وتم حساب الصدق التلازمى بين الدرجة الكلية فى البطارية ومحكات التحصيل والكفاءة فى مراكز التدريب المهنى فبلغ معامل الارتباط بينها وبين درجات المواد النظرية ٠,٢١ وبين درجات المواد العملية ٠,١٨ وكانت معاملات الارتباط بين اختبارات الورقة والقلم والمواد النظرية ٠,١٩، والارتباط بين الاختبارات العملية والمواد العملية ٠,١٠.

واستخرجت معايير الاختبارات الفرعية التى تتألف منها هذه البطارية من أداء

عينات التقنين حيث استخرج لكل اختبار معاييره المئينية ودرجاته التائية، ومما يجدر ذكره أن هذه البطارية لم يجر لها أى تقنين جديد منذ ذلك الحين، ومع ذلك لاتزال تستخدم حتى الآن فى مصلحة الكفاية الإنتاجية والتدريب المهنى، وهو أمر يشكك فى جدواها فى اتخاذ قرارات الانتقاء بعد مرور ما يقرب من ٢٠ عاما على آخر دراسة منظمة لها.

بطارية الإستعدادات الفارقة لمورسبى:

أعد هذه البطارية فى الأصل العالم الإنجليزى ج. و. مورسبى وأصدرتها هيئة البحوث التربوية فى انجلترا وويلز (NFER) وتتألف من الاختبارات الفرعية الآتية:

١ – اختبار المسلسلات المركبة: وفكرته أشبه باختبارات المصفوفات المتتابعة لرافن.

٢ – اختبارات القدرة العامة اللفظية والعددية والإدراكية.

٣ – اختبارات السرعة وتتألف من ٦ اختبارات لقياس القدرة الكتابية والمرونة العقلية والطلاقة اللفظية والطلاقة الفكرية والسرعة اليدوية والمهارة اليدوية.

٤ – اختبار الأشكال.

٥ – اختبار القدرة الميكانيكية.

وقد نقل هذه الاختبارات إلى اللغة العربية وقننها على البيئة المصرية مصطفى الشرقاوى (٣٧) كما قنن هذه الاختبارات أيضاً على البيئة الجزائرية والليبية.

وكانت العينات المصرية تتألف من طلبة مراكز التدريب المهنى بالقاهرة، أعاد عليهم تطبيق البطارية وبلغت معاملات الثبات لاختبار المسلسلات المركبة ٠,٨٥، والقدرة العامة اللفظية ٠,٦٧ والقدرة العامة العددية ٠,٧٥ والقدرة العامة الإدراكية ٠,٧١ أما اختبارات السرعة فتراوحت معاملات ثباتها بين ٠,٦٩ للطلاقة الفكرية، ٠,٩٧ للطلاقة اللفظية، واختبار الأشكال ٠,٨١ واختبار القدرة الميكانيكية ٠,٩٣.

وحسب الصدق لاختبارات البطارية باستخدام طريقة المقارنة الطرفية وتبين أن الاختبارات تميز بين التلاميذ الممتازين والضعاف فى دراستهم بمراكز التدريب المهنى فى بعض الحرف، وحسبت المعايير من أداء عينة تتألف من طلاب المدارس

الثانوية العامة بالقاهرة وتلاميذ دبلوم مراكز التدريب المهنى بالقاهرة وطلاب السنة الأولى بكلية الآداب والعلوم الاجتماعية بجامعة وهران بالجزائر، وطلاب المدارس الثانوية بمكتب تعليم طرابلس بليبيا تمتد أعمارهم بين ١٧ – ١٨ سنة، واستخدم الباحث معيار نسبة الذكاء الانحرافية فى معظم الاختبارات.

بطارية الاختبارات المعرفية مرجعية العوامل:

هذه البطارية من إعداد هيئة خدمات الاختبارات التربوية (ETS) فى برنستون بالولايات المتحدة الأمريكية وتسمى بالإنجليزية Kit of factor-referenced cog-nitive tests وظهرت منها عدة طبعات آخرها ما صدر عام ١٩٧٦، وقد نقل هذه البطارية إلى اللغة العربية فريقان من الباحثين: الأول يتكون من سليمان الخضرى وأنور الشرقاوى ونادية عبد السلام بجامعة عين شمس) والثانى على حسين بدارى وأنور رياض عبد الرحيم (بجامعة المنيا) ولم يتعد الأمر فى حدود علم مؤلف هذا الفصل - حدود الترجمة أو التعريب، ولا تتوافر لدينا أى معلومات سيكومترية من البطارية أو مكوناتها.

تقويم بطاريات الاستعدادات المتعددة:

من الواضح أن بطاريات الاستعدادات المتعددة - التى تناولنا نماذج منها فى هذا الفصل - تختلف فى الأسس العامة عن اختبارات الذكاء العام. فاختبارات الذكاء العام تتضمن عينات من النشاط العقلى التى لها أهميتها فى المواقف التى تتلاءم معها هذه الاختبارات، أما اختبارات الاستعدادات المتعددة فإنها تسعى إلى قياس العمليات العقلية المحدودة النطاق والتى تعد ضرورية فى مقررات دراسية معينة أو فى أنماط معينة من المهن، وقد أدى اهتمام بعض علماء النفس «بالتكوين العاملى البسيط» إلى استخدام الاختبارات التى تتكون من أسئلة محدودة القيمة فى قياس النشاط العقلى المعقد الضرورى فى عمليات التعلم المدرسى وفى النشاط المهنى، وللتغلب على هذه الصعوبة لجأ بعض صناع بطاريات الاستعدادات المتعددة إلى جمع درجات الاختبارات الفرعية التى تتألف منها هذه البطاريات بطرق مختلفة ولمختلف الأغراض، وفى هذا نوع من الاعتراف بأن العوامل النوعية أو القدرات الخاصة لا تعمل منفردة، وإنما تنتظم مع غيرها من القدرات فى نمط من النشاط العقلى على درجة من التعقد النسبى.

وتؤكد نتائج البحوث أنه عند استخدام الاختبارات فى الأغراض التطبيقية - كالتوجيه والانتقاء والتصنيف فى مؤسسة معينة ولمهن خاصة - فإن الاختبارات التى تعد خصيصاً لذلك تتفوق على الاختبارات المقننة والمعدة للاستخدام العام، ويصدق هذا أيضًا على الميدان التربوى، ووراء هذا الرأى ذلك الاهتمام المتزايد فى الوقت الحاضر بالاختبارات التحصيلية، وكذلك فإن الاختبارات التى تقنن على عينات متنوعة من فئات مهنية وتعليمية مختلفة تفيد أكثر من غيرها فى هذه الأغراض العملية، وهذا ما حدث بالفعل فى تقنين معظم بطاريات الاستعدادات المتعددة.

وحينما يستخدم الأخصائى النفسى بطاريات الاستعدادات المتعددة فى التمييز بين الأفراد فى مدى صلاحيتهم لأنواع التعليم والمهن المختلفة فعليه أن يتيقن من الصدق الفارق differential validity للبطاريات قبل استخدامها فقد يجد أن بعض الاختبارات الفرعية أقدر على التنبؤ بنتائج معيّنة من غيرها وهنا يقع صانعو هذه البطاريات فى خطأ فادح حين يؤكدون فى تقنينها الفروق بين متوسطات درجات الاختبارات الفرعية التى تتألف منها بالنسبة للجماعات المهنية والتعليمية المختلفة، ويهملون ما يمكن أن يوجد بين درجات هذه الجماعات من تداخل، فبين طلاب الطب والهندسة مثلا على الرغم من اختلافهما كثير من العناصر المشتركة، ومن هنا يتوافر لهم أنماط متشابهة من اختبارات العوامل المتعددة، إلا أن بينهم أيضًا فروق فى السمات العقلية وغير العقلية التى لا تقيسها هذه البطاريات.

الفصل العاشر

اختبارات القدرات المتخصصة

أشرنا فيما سبق إلى أن اختبارات الذكاء العام لا تشمل فيما تقيسه بعض «القدرات العقلية» وقد ظهر هذا النقص حتى قبل الاهتمام ببطاريات الاستعدادات المتعددة، وتمثل ذلك فى ظهور الاختبارات التى تقيس ما كان يسمى «الاستعدادات الخاصة»، وأشهرها تبكيرا اختبارات «الاستعداد الميكانيكى» ثم تزايدت الحاجة إلى هذه الاختبارات التى تقيس ما يمكن أن نسميه «القدرات المتخصصة» Specialized abilities بزيادة الاهتمام فى علم النفس التطبيقى بمشكلات التوجيه والانتقاء والتوزيع والإرشاد النفسى، فتكاثر عددها وشملت مجالات أكثر تنوعا ومن ذلك ميادين النشاط الموسيقى والفنى والمهنى والكتابى، بالإضافة إلى النشاط الميكانيكى الذى أشرنا إليه.

وقبل أن نتناول نماذج من اختبارا القدرات المتخصصة بشىء من التفصيل لابد أن نوضح مكانة هذه الفئة من الاختبارات فى النظرية الحديثة للقدرات العقلية، وبإيجاز نقول أن هذه الاختبارات ظهرت فى وقت كان الاهتمام كله مركزا على ما يسمى «اختبارات الذكاء العام» بهدف أن تكون مكملة لمفهوم «نسبة الذكاء» فى وصف النشاط العقلى للمفحوص، إلا أنه مع التقدم فى مناهج البحث وفى نظرية القدرات العقلية، وخاصة بعد شيوع استخدام منهج التحليل العاملى والنماذج المؤسسة عليه (٢٤) أصبح من الشائع اعتبار مفهوم «الذكاء» ذاته يتألف من عدد كبير من «القدرات العقلية» يصل فى نموذج جيلفورد (٢٤، ٥١) إلى أكثر من ١٢٠ قدرة ويزيد عن ذلك كثيرا فى النماذج المعاصرة لتجهيز المعلومات (٢٤) وزاد نتيجة لذلك الاهتمام ببطاريات الاستعدادات المتعددة، بل إن بعض الاختبارات التقليدية «للاستعدادات الخاصة» أصبحت من مكونات بعض بطاريات الاستعدادات كما أشرنا فى الفصل السابق.

والسؤال الذى تطرحه أنستازى (٤٠) هو: ما هو موضع هذه الاختبارات فى الوقت الحاضر؟ الواقع أن لهذه الاختبارات وظيفتين مهمتين فى التقويم النفسى المعاصر هما:

١ – توجد بعض مجالات السلوك الإنسانى لا تتوافر لها مقاييس جيدة فى بطاريات الاستعدادات المتعددة ومن ذلك الإدراك البصرى والسمعى والمهارة الحركية

والمواهب الموسيقية والفنية والسلوك الابتكارى، وقد يكون السبب فى ذلك أن المواقف التى تتطلب استخدام هذه الاختبارات على درجة كبيرة من النوعية بحيث لا تبرر إدماجها فى بطاريات الاستعدادات المتعددة، ومعنى ذلك أن هذه الاختبارات ستظل فئة متميزة من الاختبارات العقلية حتى تتوافر بطاريات «شاملة» لقدرات الإنسان، وهذا ما نتنبأ بحدوثه فى المستقبل القريب لعلم القياس النفسى.

٣ – تستخدم اختبارات القدرات المتخصصة حتى فى الأحوال التى يوجد لها ما يقابلها فى بطاريات الاستعدادات المتعددة، ومن ذلك اختبارات القدرات الميكانيكية أو الكتابية، والسبب فى ذلك أن بعض اختبارات القدرات المتخصصة تتوافر لها بيانات تقنين وافية، بالإضافة إلى المرونة فى استخدام الاختبارات، وشمول الميدان الذى تقيسه، فبدلا من استخدام اختبار واحد للقدرة الميكانيكية فى إحدى البطاريات المتعددة الاختبارات، يمكن للباحث أن يطبق بطارية لاختبارات القدرة الميكانيكية إذا تطلب الأمر قياس مختلف جوانب هذه القدرة فى المفحوصين.

ويتوافر فى الوقت الحاضر عدد هائل من الاختبارات التى تقيس «القدرات المتخصصة» ابتداء من المستوى الحسى (السمعى والبصرى) حتى مستوى التفكير الابتكارى والنشاط المهنى. ولايتسع المقام لتناول هذه الاختبارات بأى نوع من العرض الشامل ولذلك سوف نقتصر على طائفة من هذه القدرات وما يقيسها من اختبارات.

اختبارات القدرات الحركية:

ظهرت الاختبارات التى تقيس خصائص الاستجابات الحركة كالسرعة والتآزر منذ فترة طويلة، ويهتم معظمها بالمهارة اليدوية Manual dexterity وقليل منها يهتم بحركات الساق والقدم والتى تتطلبها بعض المهن والأعمال، ويقيس بعضها مزيجا من الاستعدادات الحركية والإدراكية والمكانية والميكانيكية، وفى أغلب الأحوال فإن اختبارات القدرات الحركية تتطلب أجهزة فى قياسها رغم ظهور بعض اختبارات الورقة والقلم فى أغراض التقويم النفسى الجماعى، وقد أشرنا فى الفصل الخامس إلى بطاريتى فلاناجان والاستعدادات العامة اللتين تضمنتا بعض الاختبارات الحركية الجماعية إلا أن الدراسات التى قام بها فلشمان (٤٨) وملتون (٥٤) تؤكد عدم وجود علاقة بين اختبارات الورقة والقلم واختبارات الأجهزة التى تقيس نفس القدرات

الحركية، وفيما يلى أمثلة للأجهزة التى تستخدم فى قياس النشاط الحركى:

١ – مقياس القوة العضلية (الديناموميتر) Dynamometer وتستخدم فى قياس قوة قبضة اليد.

٢ – جهاز قياس زمن الرجع .

٣ – اختبار Crawford Small Parts Dexterity ويقيس المهارة اليدوية البسيطة.

٤ – اختبار Purdue Pegboard ويقيس النشاط الحركى الدقيق والنشاط الحركى الغليظ.

٥ – اختبار Stromberg Dexterity ويقيس حركات اليد والذراع.

اختبارات القدرات الميكانيكية:

إن ما يسمى القدرة الميكانيكية ليس قدرة موحدة لا تقبل التحليل، وإنما هى فى الواقع مركب من القدرات الحسية والإدراكية والحركية، بالإضافة إلى قدرات إدراك العلاقات المكانية، واكتساب المعلومات الميكانيكية، وفهم العلاقات الميكانيكية.

ولذلك فإن الاختبارات التى تقيس القدرات الميكانيكية إنما تهتم بمستوى من الأداء أعلى مما تتضمنه الاختبارات الحسية والإدراكية (التمييز السمعى والبصرى والحركى والتآزر العضلى والمهارة اليدوية).

ويعد اختبار التجميع assembly الذى أعده ستنكويست Stenquist عام ١٩٢٣ أول اختبار ظهر فى هذا الميدان يقيس قدرة المفحوص على تركيب أجزاء الأجهزة الميكانيكية مثل جرس الدراجة، أو قفل، أو مصيدة فئران، ويتألف من ٣ سلاسل تستخدم مع مستوى عمرى يمتد بين الطفولة والرشد.

وقد عدل هذا الاختبارا فى جامعة مينيسوتا عام ١٩٣٠ وأصبع يطلق عليه اختبار مينيسوتا للتجميع الميكانيكى Minnesota Mechanical Assembly Test وأضيفت إليه أجهزة جديدة، ويصحح هذا الاختبار كسابقه فى ضوء معدل الاستجابة ودقتها، وقد أكدت الدراسات التى أجريت عليه أنه على درجة كافية من التنبؤ بنجاح طلاب التعليم الفنى المبتدئين فى أعمال الورشة. كما ارتبط بالنجاح فى بعض المهن الميكانيكية.

وفى عام ١٩٣٠ ظهر اختبارا آخر فى جامعة ميلسوتا هو اختبار مينيسوتا للعلاقات المكانية Minnesota Spatial Relations Test ويتألف من سلسلة من ٤ لوحات خشبية يتكون كل منها من ٥٧ قطعة خشبية مختلفة الأشكال، وبعضها غير مألوف ويطلب من المفحوص أن يضع هذه القطع فى أماكنها فى فتحات اللوحة الخشبية، وفى رأى كثير من النقاد أن هذا الاختبار يقيس السرعة والدقة فى الاستجابة لتفاصيل العلاقات المكانية، وقدرة المفحوص على التعامل مع الأشياء والمواد المحسوسة، وبعبارة أخرى فإنه لايقيس القدرة على حل المشكلات ذات الطبيعة الميكانيكية، كما لايصلح لقياس القدرة على معالجة الأشياء الصغيرة بدقة.

وفى عام ١٩٤٨ ظهر اختبار مينيسوتا للوحات الأشكال الورقية Revised Minnesota Paper Formboard. ويعرض على المفحوص على ورقة مطبوعة نفس المشكلات التى كانت تعرض عليه فى لوحات الأشكال الحقيقية، ففى كل سؤال من أسئلة هذا الاختبار يعرض على المفحوص جزءان أو أكثر من شكل هندسى بحيث لو جمعت معا بطريقة صحيحة فإنها تكون الشكل الكلى الكامل، وعلى المفحوص تحديد الشكل الذى يمكن أن يجمع مع الأجزاء الأخرى من بين ٥ بدائل اختبارية، وهكذا يقيس هذا الاختبار القدرة المكانية كما تحددت فى البحوث العاملية للقدرات العقلية. وتؤكد نتائج البحوث التى أجريت على هذا الاختبار أنه يرتبط ارتباطا دالا بنوع الأداء الميكانيكى والنجاح فى الرسم الهندسى والهندسة الوصفية، واستخدمت فى حساب صدقه محكات مختلفة مثل درجات مقررات الهندسة والمواد الفنية والورش الميكانيكية وتقديرات المشرفين ومقاييس الكفاءة الإنتاجية، وبوجه عام يمكن القول أن طلاب الهندسة والعاملين بالمهن الميكانيكية يحصلون على درجات فى هذا الاختبار أعلى من غيرهم، وقام بتقنينه على البيئة المصرية فؤاد أبو حطب وآمال صادق (٢٦).

وتوجد فئة أخرى من اختبارات القدرة الميكانيكية تقيس المعلومات الميكانيكية والاستدلال الميكانيكى والفهم الميكانيكى، وهذه جميعا تتطلب من المفحوص ألفة بالآلات الشائعة وبالعلاقات الميكانيكية، إلا أنها لا تفترض فى المفحوص توافر معلومات فنية أكثر مما يمكن أن توفره الحياة اليومية فى المجتمعات الصناعية، ومن أشهر هذه الاختبارات تلك التى أعدها بينيت وزملاؤه (٤٤) والتى تسمى اختبارات الفهم الميكانيكى Tests of Mechnical Comprehension.

وتعرض على المفحوص مشكلات ميكانيكية مصورة وعدة حلول لكل منها، وعليه أن يختار الحل الصحيح، ولهذه الاختبارات ثلاثة مستويات من الصعوبة، وتهدف إلى قياس القدرة على فهم المبادئ الفيزيائية والميكانيكية فى مواقف بسيطة نسبيا.

ومن الاختبارات الأحدث نسبيا فى قياس القدرة الميكانيكية اختبار الدافعية الميكانيكية المسمى Mellanbruch Mechanical Motivation Tests وهو اختبار مصور يتطلب من المفحوص التعرف على أشياء معينة شائعة فى المجتمعات الصناعية المتقدمة، والاختبار الثانى هو SRA Mechanical Aptitudes Test ويتألف من ٣ اختبارات فرعية تقيس:

(أ) المعلومات الميكانيكية.
(ب) إدراك الأشكال والتصور البصرى المكانى.
(جـ) حل المشكلات الحسابية الخاصة بالأعمال الميكانيكية (فى صورة جداول ورسوم بيانية).

ويجب أن نؤكد أن هذه الاختبارات التى تقيس الفهم الميكانيكى تتطلب تعديلا شاملا عند محاولة استخدامها فى البيئة المصرية، وذلك لأنها مشبعة تشبعا كبيراً بالعناصر الثقافية السائدة فى المجتمعات التى صنعت فيها.

اختبارات القدرة الكتابية:

القدرة الكتابية Clerical ability ـ كالقدرة الميكانيكية ـ ليست قدرة بسيطة، وإنما هى على درجة من التعقد والتركيب ، ورغم أن بعض بطاريات الاستعدادات المتعددة التى تناولناها فى الفصل الخامس ـ وخاصة بطارية الاستعدادات الفارقة وبطارية الاستعدادات العامة ـ تضمنت اختبارات فرعية لقياس القدرة الكتابية فإن تراث التقويم النفسى يتضمن طائفة من الاختبارات التى تقيس هذه القدرة على وجه الخصوص، بل توجد بطاريات لقياس القدرة الكتابية، وعموما نستطيع القول أن اختبارات هذه القدرة تجمعها خصائص مشتركة يوضحها الجدول الآتى الذى يلخص محتوى ٦ بطاريات للقدرةالكتابية واختباراتها الفرعية، ومنه يتضح أن البطاريات الشاملة تسعى إلى قياس عدد من العمليات النوعية، بينما تسعى البطاريات الأقل

شمولا إلى قياس السرعة والدقة الإدراكيتين والتى تعد أحد جوانب القدرة الكتابية وليست كلها، وقد يكون من العوامل الكامنة وراء قصر القدرة الكتابية على السرعة الإدراكية ما تؤكده نتائج البحوث من أهمية القدرات الإدراكية بوجه عام فيها، إلا أن «القدرة الكتابية» تشمل أنواعا متنوعة من الأعمال تختلف فى مدى القدرة ومستواها، ولذلك فإن مقاييسها تضمن عمليات عقلية أخرى أساسية، بالإضافة إلى بعض القدرات الحسية والمهارات اليدوية عند الضرورة، ويلخص الجدول رقم (٢٩) الاختبارات الفرعية التى تتضمنها ٦ بطاريات للقدرة الكتابية.

جدول رقم (٢٩)

يلخص الاختبارات الفرعية التى تتضمنها ٦ بطاريات للقدرة الكتابية

البطارية	الاختبارات الفرعية
بطارية ديترويت Detroit	الخط: السرعة والدقة المراجعة: السرعة والدقة الحساب البسيط السرعة والدقة فى الأداء الحركى معرفة المصطلحات التجارية البسيطة ترتيب الصور التصنيف : السرعة والدقة الترتيب الأبجدى
بطارية الاستعدادات الكتابية العامة General Clerical	المزاوجة: اكتشاف الأخطاء فى الأسماء والأعداد الترتيب الأبجدى الحساب: الجمع البسيط الحساب: تحديد أخطاء الجمع المسائل الحسابية البسيطة التهجى فهم القراءة معرفة معانى الكلمات الاستخدام اللغوى: القواعد النحوية

تابع الجدول رقم (٢٩)

جدول يلخص الاختبارات الفرعية التى تتضمنها ٦ بطاريات للقدرة الكتابية

البطارية	الاختبارات الفرعية
بطارية مينيسوتا	مقارنة الأعداد مقارنة الأسماء
بطارية بردو Purdue	التهجى الحساب المراجعة: السرعة معرفة معانى الكلمات النقل والنسخ: الدقة الاستدلال
بطارية التوظيف Short Employment	العمليات العددية معانى الكلمات التصنيف
بطارية تورس Turse	المهارات اللفظية المهارات العددية التعليمات المكتوبة المراجعة: السرعة التصنيف الترتيب الأبجدى

وتوجد فى المكتبة العربية بطاريتان لقياس القدرة الكتابية هما:

(١) **اختبارات القدرة الكتابية**: وهى من إعداد محمد عبدالسلام أحمد (٤٦) عام ١٩٦٠ وتتألف من اختبارين هما: اختبار تصنيف الأعداد، واختبار تصنيف الأسماء، اعتمدا على تحليل للأعمال الكتابية قام به الدكتور محمد عبدالسلام أحمد بمعاونة ديوان الموظفين عام ١٩٥٣ ، وقنن الاختبار الأول على عينة حجمها ٨٢٢ مفحوصاً من الجنسين الحاصلين على شهادة الثانوية العامة أو دبلوم التجارة

الثانوية تتراوح أعمارهم بين ١٨ و ٢٨ سنة، وقنن الاختبار الثانى على عينة عددها ٨٨٠ فردا من الجنسين تتراوح أعمارهم أيضًا بين ١٨، ٢٨ سنة من حملة الثانوية العامة المتقدمين لشغل الوظائف الكتابية عن طريق إعلانات ديوان الموظفين فى ذلك الوقت. وأعدت للاختبارين معايير مئينية وتائية وحسب لهما معامل الثبات بطريقة التجزئة النصفية فبلغ ٠,٩٧ . (ن = ١٠٠) والواقع أن استخدام طريقة التجزئية النصفية فى حساب اختبارات القدرة الكتابية - التى تعتمد اعتمادا كليا على السرعة - يثير كثيرا من المشكلات أشرنا إليها فى الفصول السابقة إلا إذا كان المؤلف لجأ إلى تجزئة زمن الاختبار، وهذا ما لم توضحه كراسة تعليمات هذه الاختبارات (١٥ : ٣١). أما عن صدق الاختبارين فقد حسب مؤلفهما معاملات الارتباط بينهما وبين اختبارات المهن الكتابية والذكاء الاجتماعى (ن = ٤٧) من طالبات معهد السكرتارية بالعباسية) فلوحظ أن الاختبارين يشتركان مع الاختبارات التى تقيس الجوانب الاتية: السرعة. والدقة، سهولة استخدام اللغة، الألفة بالأعداد، الذاكرة المباشرة، المهارة اليدوية.

(٢) **اختبارات المهن الكتابية:** من إعداد محمد عماد الدين إسماعيل وسيد عبد الحميد مرسى (٣٣) وتقيس مكونات ثلاث هى: القدرة العددية، السرعة والدقة، الاستدلال اللغوى، وقد حسب ثبات الاختبارات الثلاث (ن = ١٠٠ من طالبات معهد السكرتارية) بطريقة التجزئة النصفية فبلغت ٠,٨٦ ، ٠,٨٤ على التوالى، ومرة أخرى نثير مسألة استخدام الطريقة الملائمة لحساب ثبات الاختبارات التى تعتمد على السرعة، واستخدم فى حساب صدق الاختبارات تقديرات مادة المحاسبة والترجمة وتقديرات المدرسين والمشرفين وتراوحت معاملات الارتباط بين ٠,٤٢ ، ٠,٤٥ وحسبت المعايير المئينية للاختبارات.

اختبارات القدرات الموسيقية:

يوجد فى الوقت الحاضر عدد من الاختبارات التى تصلح لقياس القدرات الموسيقية، قد يكون أقدمها ظهورا أو أكثرها أهمية اختبارات سيشور التى تسمى: Sea-shore Measures of Musical Talent.

وقد ظهرت فى صورتها الأصلية عام ١٩١٩، واستمرت شائعة الاستخدام فى

الولايات المتحدة دون تعديل حتى عام ١٩٣٩ حيث عدلت فى صورتها التى تستخدم حتى الآن، وقد كانت هذه الاختبارات نتيجة للبحوث الواسعة النطاق فى ميدان سيكولوجية الموسيقى والتى اجريت بجامعة أيوا تحت إشراف كارل سيشور، وتتألف هذه البطارية فى صورتها الحديثة من ٦ اختبارات هى، تمييز الأصوات pitch، شدة الصوت loudness تذكر الإيقاعات rhythm ، الزمن time، نوعية الصوت timbre، تذكر الألحان tonal memory.

ويتألف اختبار تمييز الأصوات من ٥٠ زوجا من النغمات، وعلى المفحوص أن يحدد بالنسبة لكل زوج ما إذا كانت النغمة الثانية أحد أم أغلظ من النغمة الأولى وفى اختبار شدة الصوت يطلب منه أن يحدد بالنسبة لكل زوج من ٥٠ زوجا من النغمات ما إذا كانت النغمة الثانية أشد أم أضعف من الأولى، ويطلب منه فى اختبار تذكر الإيقاعات أن يحدد ما إذا كان النمطان الإيقاعيان اللذان يتألف منهما كل سؤال (من ٣٠ سؤالا) متشابهين أو مختلفين: وفى اختبار الزمن يحدد المفحوص ما إذا كانت النغمة الثانية أطول أم أقصر من النغمة الأولى (٥٠ سؤالا)، وفى اختبار نوعية الصوت يقوم المفحوص بالحكم على النغمة الثانية بأنها متشابهة أو مختلفة عن النغمة الأولى من حيث نوعية الصوت (البناء الهارمونى)، ويتكون هذا الاختبار من ٥٠ سؤالا، أما الاختبار الأخير، وهو تذكر الألحان، فيتكون من ٣٠ زوجا من النماذج اللحنية، وفى كل زوج توجد نغمة واحدة مختلفة فى النموذج الثانى، وعلى المفحوص أن يحدد أى النغمات اختلفت ويحدد ترتيبها الرقمى.

وتصلح هذه الاختبارات للتطبيق على المفحوصين ابتداء من الصف الرابع الابتدائى حتى أعلى المستويات التى تتطلب إعدادا موسيقيا رفيعا للمراهقين والراشدين.

وقد ظهرت هذه الاختبارات فى أصلها الأجنبى مسجلة على اسطوانات، وقد أعدتها باللغة العربية آمال أحمد مختار صادق (٥) مسجلة على شرائط*، وأعدت تعليمات تطبيقها، كما صممت كراسة إجابة مختلفة عن الطبعة الأجنبية، وحسبت معاملات الثبات بطريقة التجزئة النصفية وتصحيح سبيرمان براون على عينة من

* وظهرت طبعة ٢٠٠٧ علي C.D.

١٦٣ طالبا وطالبة من حملة الثانوية العامة المتقدمين للالتحاق بالمعهد العالى للتربية الموسيقية وتراوحت معاملات الثبات بين ٠,٥٨ (اختبارالزمن)، ٠,٨٨ (اختبار تذكر الألحان). أما عن صدق هذه الاختبارات فقد حسبت الباحثة معاملات الارتباط بين الاختبارات الفرعية التى تتألف منها البطارية لعينات من طلاب دور المعلمين والمعلمات (ن = ١٧١(وطلاب المعاهد الموسيقية العليا (ن = ٦٠)، وتلاميذ المدارس الإعدادية العامة (ن = ١١١)، وتلاميذ المعهد القومى للموسيقى (ن = ٧٤)، كما حسبت معاملات الارتباط بين اختبارات سيشور واختبارات القدرات الموسيقية للأطفال (بنتلى) لعينة تلاميذ المرحلة الإعدادية العامة. وتلاميذ المعهد القومى للموسيقى، وكذلك معاملات ارتباطها باختبارات ونج للذكاء الموسيقى، وأريجون للتذوق الموسيقى، وفارنم للتدوين الموسيقى وكاتل للذكاء العام، واختبار الذكاء المصور الذى اعده أحمد زكى صالح، وتتوافر فى كراسة التعليمات بيانات كافية عن المعالم السيكومترية لهذه البطارية. أما عن معايير هذه البطارية فقد حسبت المئينيات الخاصة بكل اختبار فرعى ولكل عينات فرعية من عينة التقنين (٥) كما تأكد الصدق التنبؤى لبطارية سيشور من خلال دراسة تتبعية قامت بها آمال صادق، وإكرام مطر (١٠) استخدم فيها أسلوب تحليل الانحدار المتعدد وتأكد فيها ارتفاع القيمة التنبؤية لاختبارات سيشور فى مجال القبول للتعليم الموسيقى فى مصر بمقارنتها بالأساليب التقليدية التى تستخدم حتى وقتنا الحاضر فى انتقاء الطلاب الجدد فى كلية التربية الموسيقية جامعة حلوان.

وقد صدرت من هذه الاختبارات ثلاثة تقارير تضمنت المعالم السيكومترية لعينات مختلفة. ولمزيد من التفصيل راجع كراسة التعليمات ٢٠٠٦ .

وقد أعدت نفس الباحثة (٦) صورة مختصرة لبطارية سيشور ووصفت معالمها السيكومترية بوضوح، كما استخدمت فى بحوث متعددة لطلاب الماجستير والدكتوراة (٢٠٠٦) فقد قام خالد يحى مندوه (٢٠٠٦) فى رسالته للدكتوراه بحساب معامل الإتساق الداخلى لبطارية سيشور المختصرة على طلاب وطالبات كلية التربية النوعية – جامعة عين شمس (ن =٥٠) وحصل على معاملات دالة عند مستوى ا. و تراوحت بين ٠,٤١ و٠,٦٦ كما تم حساب ثبات البطارية المختصرة باستخدام معادلة كيودر – ريتشاردسون وتراوحت معاملات الثبات بين ٠,٣٦، ٠,٦٣ وبطريقة التجزئة النصفية

بين ٠,٤٠ ، ٠,٩٨ وجميع المعاملات المحسوبة دالة عند مستوى ٠,٠١ ، كما تم تحديد معاملات الثبات على عينة أخرى قوامها ٧٨ مفحوصا باستخدام التجزئة النصفية، وتطبيق معادلة ألفا كرونباك وطريقة جاتمان وأثبتت كل الطرق ثبات البطارية. وقد قامت معدة البطارية باستخراج الدرجات المئينية للدرجات الخام من المصدر السابق وأيضا حساب المتوسطات والانحرافات المعيارية والتى يمكن الاستفادة منها من حساب الدرجات المعيارية لاغراض المقارنة.

ويوجد اختبار آخر للقدرة الموسيقية ظهر عام ١٩٦٣ أعده فى إنجلترا أرنولد بنتلى Bentley خصيصًا لأطفال المرحلة الابتدائية. (من ٧ – ١٢ سنة) ويقيس ٤ قدرات هى: تمييز الأصوات ، تذكر النغمات، تحليل التألفات، تذكر الإيقاعات، ويمكن تطبيق هذه الاختبارات على مستويات عمرية أعلى من مراعاة معايير هذه الأعمار عند تفسير الدرجات.

ويتكون اختبار تمييز الأصوات من ٢٠ زوجا من الأصوات «النقية» يستمع إليها المفحوص ثم يطلب منه أن يحدد ما إذا كان الصوت الثانى من كل زوج متشابها مع الصوت الأول أو مختلفا عنه، وفى حالة الاختلاف على المفحوص أن يحدد ما إذا كان أعلى (أى أحد) أو أسفل (أى أغلظ) من الصوت الأول، أما اختبار تذكر النغمات فيتكون من ١٠ أسئلة يتطلب كل منها مقارنة ثنائية بين جملتين لحنيتين قصيرتين، ويتكون كل منها من ٥ نغمات، ويطلب من المفحوص أن يحدد ما إذا كانت الجملة الثانية متشابهة مع الأولى أو مختلفة عنها وإذا حدث اختلاف عليه أن يحدد ترتيب النغمة المختلفة. وفى اختبار تحليل التآلفات يطلب من المفحوص تحديد عدد النغمات التى يتكون منها كل تآلف (يتكون الاختبار من ٢٠ تألفا)، أما اختبار تذكر الإيقاعات فيتكون من ١٠ مفردات يتطلب كل منها مقارنة ثنائية بين جملتيين إيقاعيتين، وتتكون كل واحدة من ٤ وحدات إيقاعية أو ما يعادلها من النماذج الإيقاعية المختلفة، ويطلب من المفحوص تحديد ما إذا كانت الجملة الثانية فى كل سؤال متشابهة مع الجملة الأولى أو مختلفة عنها، وفى حالة الاختلاف يكون عليه تحديد رقم الوحدة المختلفة.

وقد أعدت هذه الاختبارات باللغة العربية أيضًا آمال أحمد مختار صادق عام ١٩٧٢ (٧)، وظهرت اختبارات القدرات الموسيقية للأطفال مسجلة على شرائط (بدلا

من الاسطوانات فى الأصل الأجنبى) وأعدت لها كراسة إجابة خاصة، كما فصلت تعليمات تطبيقها، وقد استخدم المؤلف الأصلى فى حساب ثبات هذه الاختبارات طريقة إعادة الاختبار، وبلغ معامل الارتباط ٠,٨٤، واستخدمت الباحثة المصرية هذه الاختبارات فى دراستين عامليتين للقدرة الموسيقية (٩، ٦١)، إحداهما على عينة من تلاميذ المرحلة الإعدادية (ن = ١١١)، والأخرى على عينة تناظرها فى العمر الزمنى من تلاميذ المعهد القومى للموسيقى (الكونسيرفتوار) عددها ٧٤ تلميذا. وتراوحت اشتراكيات الاختبارات الفرعية بين ٠,٤٢ (لاختبار تذكر الإيقاعات لعينة تلاميذ المرحلة الإعدادية العامة)، ٠,٦٨ (لاختبار تمييز الأصوات لنفس العينة)، أما اشتراكيات الدرجة الكلية للبطارية فبلغت ٠,٨٨ فى الحالتين.

ولحساب صدق الاختبارات حسبت معاملات الارتباط بين الاختبارات الفرعية، ومعاملات ارتباطها باختبارات سيشور التى أشرنا إليها، وباختبارات كاتل للذكاء العام، وكذلك حسبت المعايير المئينية لعينتين من التلاميذ: إحداهما مجموعة ٩ – ١٢ سنة (ن = ١٦٢)، والأخرى مجموعة ١٢ – ١٥ سنة (ن = ٢٦٣).

وقد صدر التقنين الثالث لهذه الاختبارات عام ٢٠٠٧ (٧).

ويوجد اختبار آخر ظهر فى إنجلترا عام ١٩٤١ ثم ظهرت طبعته الجديدة المعدلة عام ١٩٥٨ ثم عام ١٩٦٢ وهو اختبار ونج للذكاء الموسيقى Wing Standardized Tests of Musical Intelligence *.

ويصلح للتطبيق ابتداء من سن ٨ سنوات،ويختلف عن اختبار سيشور فى أنه لايؤكد الجانب «الذرى» التحليلى كما يتمثل فى عناصر الإحساس البسيط السائد فى اختبارات سيشور، ولذلك فإنه يعتمد على محتوى موسيقى واقعى حيث تستخدم موسيقى البيانو فى كل اختبار من الاختبارات السبع التى تتألف منها بطارية ونج وهى: تحليل التآلفات، تمييز الأصوات، ذاكرة الأصوات، التوافق الموسيقى (الهارمونى)، الشدة، الإيقاع لجمل الموسيقية، وتتطلب الاختبارات الثلاث الأولى تمييزا حسيا على مستوى أكثر تعقدا من اختبارات سيشور، أما فى الاختبارات الأربع

---

* نشرت اجراءات تحديث تقنين اختبارات القدرات الموسيقية فى مصر فى كتاب: دراسات وبحوث فى سيكولوجية الموسيقى والتربية الموسيقية الطبعة الثانية، الانجلو المصرية ٢٠٠٧ .

الأخرى فإن المفحوص يقارن النوعية الجمالية فى مقطوعتين موسيقيتين، ومعنى ذلك أن بطارية ونج تؤكد جانب التذوق الموسيقى، ويقترح ونج استخدام الاختبارات الثلاث الأولى كبطارية مختصرة مع الأطفال الصغار أو كأداة تصفية مبدئية.

وقد نشرت آمال أحمد مختار صادق (٣) فى عام ١٩٧٩ تقريرها عن تقنين هذا الاختبار على البيئة المصرية، وقد بلغت عينة التقنين ٣٤١ مفوحصًا من الذكور والإناث اشتقوا من مستويات تعليمية موسيقية مختلفة من دور المعلمين والمعلمات والمعهد القومى للموسيقى (الكونسيرفتوار) وكلية التربية الموسيقية بجامعة حلوان - وقد سارت الباحثة فى تقدير صدق المحتوى من خلال استطلاع رأى المهتمين بالموسيقى بدرجة «الموسيقية» فى هذه الاختبارات، أما بالنسبة لصدق التكوين الفرضى بحساب معاملات الارتباط بين الاختبارات الفرعية للبطارية وتحليل الفروق من مستويات التدريب الموسيقى وتمايز المستويات التعليمية الموسيقية، كما استخدم الصدق المرتبط بالمحكات باستخدام بطارية سيشور التى سبق تقنينها من ناحية والتحصيل الموسيقى من ناحية أخرى، وكانت هذه النتائج جميعا مدعمة لصدق البطارية.

وقد نشر تقرير آخر عن تحديث تقنين بطارية ونج متضمنة بيانات حديثة (٩: ٢٠٠٦).

أما الثبات فقد حسب بطريقتين: التجزئة النصفية، كيودر - ريتشارد سون، وحسبت المعايير المئينية للاختبارات الفرعية ولمجموع الاختبارات الإدراكية ومجموع الاختبارات التذوقية والدرجة الكلية فى البطارية، كما حسبت أيضًا نسب الذكاء الانحرافية التى أطلقت عليها الباحثة تسمية «نسب الذكاء الموسيقى» ومصطلح «الذكاء الموسيقى» هو المصطلح الأثير عند ونج مؤلف هذه الاختبارات.

ونظرا لأهمية استخدام بطارية ونج للذكاء الموسيقى سواء فى قياس الجوانب الإدراكية متمثلة فى الاختبارات الثلاثة الأولى، أو الجوانب التذوقية متمثلة فى الاختبارات الأربعة الاخيرة فقد استخدمت فى العديد من الدراسات سواء من معدة الاختبارات أو طلابها فى مرحلة الدراسات العليا فى الكليات الموسيقية بمصر أو العالم العربى، ومن هذه الدراسات دراسات خالد يحيى مندوه للماجستير والدكتوراه، وأمل سليم أيضا فى رسالتى الماجستير والدكتوراه، وحددت فى هذه الدراسات جميعها

المعالم السيكومترية للبطارية. ولمزيد من التفصيل راجع تقرير التقنين فى (٩: ٢٠٠٦).

وفى عام ١٩٧٩ نشرت آمال صادق أيضًا دراسة مهمة تتضمن عرضًا مفصلا لبطارية جديدة لقياس القدرات الموسيقية قامت بإنشائها (٤) فى محاولة منها للربط بين مجال قياس القدرة الموسيقية والتقدم فى مجال دراسة القدرات العقلية بصفة عامة وما أحدثته نتائج البحوث فيها من تغير لم يظهر له أثر فى بطاريات سيشور وونج وبنتلى التى سبق الإشارة إليها. وتشمل الاختبارات الجديدة التى قامت الباحثة ببنائها ما يلى:

(١) **اختبار الإغلاق الموسيقى**: musical closure ويتألف من قسمين أولهما من ٢٤ مفردة فى كل منها يستمع المفحوص إلى جملة موسيقية قصيرة، وبعدها بفترة يستمع إلى جملة موسيقية أكبر، ويطلب منه تحديد ما إذا كانت الجملة الأولى متضمنة فى الجملة الثانية أم لا. وثانيهما ويتألف من ١٢ مفردة، كل منها عبارة عن جملة موسيقية اختيرت من مؤلفات للبيانو، ويعرض على المفحوص (موتيف) وعليه أن يقرر ما إذا كان موجودا فى اللحن أما لا.

(٢) **اختبار تعيين الألحان**: melody identification ويتألف من صورتين إحداهما للمراهقين والراشدين والثانية للأطفال، وتتكون صورة الكبار من قسمين أولهما يتألف من ٤٨ مفردة من الألحان فى السلالم الكبيرة والصغيرة الغربية، وثانيهما من ١٢ مفردة من الألحان التى تتضمن المسافة العربية (أى تمثيل الموسيقى العربية)، أما اختبار الأطفال فيتألف من ٣٠ مفردة كلها من نوع الموسيقى الغربية، وفى جميع الحالات يقوم المفحوص بالمقارنة بين المعزوف والمدون على هيئة نوتة موسيقية.

(٣) **اختبار التعيين الإيقاعى**: rhythmic identification وقد أعد من صورتين أيضًا، وتتألف صورة المراهقين والراشدين من ٦٠ مفردة عشرة منها فى كل ميزان من الموازين الموسيقية ويطلب من المفحوص أن يقارن نمطا إيقاعيا منقرا بنمط إيقاعى مكتوب على هيئة تدوين موسيقى، أما صورة الأطفال فتتألف من ٣٠ مفردة كلها فى الميزان الموسيقى البسيط.

(٤) **اختبارالتعرف على المسافات**: interval recognition ويتألف من صورة واحدة تصلح للأطفال لسهولته الظاهرة بالنسبة للمراهقين والراشدين، وينقسم إلى قسمين: أولهما يتكون من ٣٠ زوجا من المسافات تعزف عزفا متتابعا على البيانو وباستخدام طبقات صوتية مختلفة، ويطلب من المفحوص أن يقرر ما إذا كانت المسافتان متشابهتين أومختلفتين، وفى القسم الثانى ١٠ مفردات أكثر صعوبة تعزف عزفا تتابعيا على الكمان تتضمن مسافات عربية (تحتوى المسافة اللحنية العربية).

(٥) **اختبار تقدير المسافات**: interval estimation ويتألف من صورة واحدة للمراهقين والراشدين لصعوبته الظاهرة بالنسبة للأطفال ويتألف من ٤٠ مفردة مصنفة إلى ثلاثة أقسام تتناول على التوالى المسافات الصاعدة فى الموسيقى الغربية، والمسافات الهابطة منها ومسافات الموسيقى العربية. وفى جميع الأقسام تعزف لكل مسافة مسافة مرجعية أصلية مرة واحدة على البيانو أو الكمان وبعد فترة توقف طولها ٥ ثوان تعزف ثلاث مسافات على التوالى ويطلب من المفحوص أن يحكم على أى هذه المسافات الثلاث هى الأقرب إلى المسافات الأصلية.

وقد قننت البطارية على عينة مصرية تتألف من تلاميذ المرحلة الإعدادية العامة (ن = ١١١)، وتلاميذ المرحلتين الابتدائية والإعدادية بالمعهد القومى للموسيقى (ن = ٧٤)، وطلاب الكليات والمعاهد الموسيقية (ن = ٦٠) وكليات التربية النوعية، وحسب الثبات بطريقة كيودر- ريتشارد سون، واستخدمت فى حساب الصدق محكات تتمثل فى الاختبارات التى تقيس القدرة الموسيقية والتى سبق تقنينها على البيئة المصرية (سيشور، بنتلى، ونج) كما استخدم محك التحصيل الموسيقى، وكانت المؤشرات الإحصائية لكل من الثبات والصدق فى صالح البطارية، كما استخدمت الباحثة طريقة أخرى فى حساب الصدق تتمثل فى المقارنة بين العينات ذات المستويات المختلفة من الانتقائية من كل من الأطفال والمراهقين والراشدين وكانت النتائج فى معظم الحالات دالة لصالح المجموعات الأعلى انتقائية (تلاميذ الكونسيرفتوار من الأطفال وطلاب المعاهد والكليات الموسيقية من المراهقين والراشدين) ولمزيد من التفصيل عن هذه الاختبارات راجع المصدر (٩: ٢٠٠٦)، والتى يلاحظ فيها ان أكثر الاختبارات استخداما هى اختبارات الاغلاق الموسيقى بمستوييها.

اختبارات القدرة فى الفنون البصرية :

توجد فى الوقت الحاضر طائفة من الاختبارات التى تقيس بعض جوانب القدرة فى ميدان الفنون البصرية، ومنها اختبار Graves Design Judgment Test الذى يهدف إلى قياس قدرة المفحوص على إدراك المبادئ الأساسية للعمل الجمالى والاستجابة لها وهى مبادئ الوحدة unity والسيادة dominance والتنوع variety والتوازن balance والاستمرار continuity والتناسق symmetry والتناسب pro-portion والإيقاع rhythm ، وقد تحقق ذلك بعرض ٩٠ ثنائية وثلاثة من الرسوم المجردة، فى كل منها واحد من الرسوم يتفق مع المبادئ الثمانى بينما يتمثل أحد هذه المباديء أو بعضها فى الرسم الآخر أو الرسوم الأخرى، ويطلب من المفحوص أن يفاضل بين رسوم الثنائية أو الثلاثية ويختار أحدها على أساس التفضيل الفنى. ولم يتضمن هذا الاختبار نماذج من أعمال الفن، لأن الحكم على الصور واللوحات يتأثر بالخبرة السابقة وسمات الشخصية والانفعالات ومفهوم المفحوص عن الفن ولذلك فإن الاختبار يؤكد محض الاختبار الجمالى، أو ما يسميه أحد مؤلفى هذا الكتاب «الحساسية الجمالية» (٢٥)، وبالإضافة إلى ذلك فإن الاختبار يقيس العمليات الإدراكية المعقدة وليس العناصر الحسية البسيطة التى تتدخل فى الحكم الجمالى ، وفى هذا يتشابه مع اختبارات ونج فى الموسيقى، ومع المبادئ المتضمنة فى قياس العامل العام فى الذكاء.

والاختبار الثانى فى هذا المجال هو اختبار ماير للحكم الفنى Meier Art Judg-ment Test.

ويهدف إلى قياس الحكم الجمالى بطريقة «كلية»، ويختلف عن اختبار جريفز فى أنه يتألف من ١٠٠ زوج من اللوحات الفنية مرسومة بالأبيض والأسود. ويمثل أحد عنصرى الثنائية عملا فنيا ممتازا، والعنصر الآخر أدخل عليه بعض التغيير والتعديل فى أحد الجوانب الفنية المهمة تجعله أقل جودة من العنصر الأول. ويحدد المفحوص أى الجوانب قد تغير (الشكل أو الزوايا مثلا) إلا أنه لايعلم أى اللوحتين هى اللوحة الأصلية، ويطلب من المفحوص أن يحدد تفضيله لأى منهما.

ويرى ماير أن هذا الاختبار يقيس الحكم الجمالى ، إلا أنه فى رأى أحد مؤلفى هذا الكتاب يقيس «الحساسية الجمالية» أيضًا (٢٥ : ٥)، والفرق بينه وبين اختبار جريفز فرق فى المحتوى وليس فى العمليات الأساسية المتضمنة.

وإذا كان اختبار جريفز وماير ينتميان إلى ميدان التذوق الفنى فإن اختبارى كنوير Knauber وهورن Horn ينتميان إلى ميدان الابتكار الفنى ، ويسمى ألهما Knauber Art Ability Test ويصلح للمراهقين والراشدين ويتطلب من المفحوص ما يأتى:

١ – أن يرسم من الذاكرة أشكالا فى حدود المكان المسموح به.

٢ – أن يرسم شخصيات نمطية (بابا نويل مثلا).

٣ – ترتيب مؤلف معين فى حدود المكان المسموح به.

٤ – ابتكار وتكملة رسوم معينة من عناصر معطاة.

٥ – تحديد الأخطاء فى أعمال مرسومة (مثل المنظور الخاطئ).

٦ – إنتاج أعمال فنية تدل على الخيال والحذق والتمثيل الرمزى.

أما اختبار هورن ويسمى Horn Art Aptitude Inventory فقد أعد خصيصاً للاستخدام فى أغراض قبول طلاب كليات الفنون الجميلة، ويقيس على وجه الخصوص الابتكارية فى الفن مستخدما فى كل سؤال وصفا بسيطا مرنا أو مرشدا لكل رسم. ويتألف الاختبار من قسمين: الأول يتطلب من المفحوص أن يرسم ٢٠ شيئا مألوفا (منزل- كتاب... إلخ) فى وقت قصير، كما يتطلب منه إنتاج رسوم تجريدية بسيطة مبتدئا بمجموعة من المثلثات والمربعات، وغيرها. أما القسم الثانى فيتطلب مستوى أعلى من الابتكارية وفيه يعطى للمفحوص ٢٠ بطاقة مستطيلة على كل منها خطوط تفيد كبداية لصورة ينتجها المفحوص، ويحكم على الرسوم على أساس الخيال الابتكارى من ناحية والخصائص الفنية من ناحية أخرى.

ومن الاختبارات المصرية التى تقيس القدرات فى الفنون البصرية اختبارات تكميل الأشكال وتكميل الصور والتقدير الجمالى التى أعدها الدكتور محمد عماد الدين إسماعيل (٣٢) لقياس عاملى الطلاقة والتقدير الجمالى، ويقيس اختبار تكميل الأشكال عامل الطلاقة وفيه يطلب من المفحوص أن يضيف إلى خطوط مرسومة بضعة خطوط أخرى بحيث تجعل منه صورة لشىء مألوف، ويقيس اختبار تكميل الصور عامل الطلاقة أيضاً وفيه يطلب من المفحوص أن يكتب فى مكان معين من صورة

مرسومة أكبر عدد من الأشياء التى يمكن رسمها، أما اختبار التقدير الجمالى فهو تقنين مصرى لاختبار ماير الذى أشرنا إليه، وقد حسب صدق الاختبارات بإيجاد معاملات الارتباط بين كل منها وبين المحك (وهو تقدير المدرسين للقدرة الفنية) فكانت النتاج كالآتى: ٠,٤٤ لاختبار تكميل الأشكل، ٠,٣٤ لاختبار تكميل الصور. ٠,٤٤ لاختبار التقدير الجمالى (ماير). أما عن الثبات فقد حسب بمعادلة كيودر ـ ريتشارد سون وكان معامل ثبات اختبار تكميل الأشكال ٠,٥٥ واختبار تكميل الصور ٠,٨٣ واختبار (ماير) للتقدير الجمالى ٠,٥٤.

وحسبت معايير الدرجات التائية للاختبارات الثلاثة، وكانت عينة التقنين فى كل الأحوال ٦٥ طالبا من المعهد العالى للتربية الفنية.

اختبارات التفكير الابتكارى:

من أهم خصائص التقويم النفسى المعاصر زيادة الاهتمام بقياس الابتكارية مسايرة للنشاط المتزايد فى مختلف ميادين علم النفس لدراسة طبيعة النشاط الابتكارى والعوامل المؤثرة فيه، ولايتسع المقام للإفاضة فى هذه المسائل، ويمكن الرجوع إلى بعض المصادر المتاحة فى اللغة العربية (٢٤) ونستطيع القول بإيجاز أن الاهتمام بالتفكير الابتكارى يعبر عن حاجة المجتمعات المعاصرة لزيادة وتنمية ثروتها البشرية من العلماء والمهندسين والقادة الإداريين والفنانين، ويرى كثير من العلماء المهتمين بهذا الميدان أن الابتكار عملية أساسية فى هذه الأحوال جميعا، فهو كما يقول تايلر «يختلف فى العمق والمجال وليس فى النوع» ومن رأيه عدم التمييز بين الإبتكار العلمى والابتكار الفنى لأن الابتكار «يتضمن نظرة للمشكلات أكثر أساسية من التدريب المهنى».

ويعود الفضل إلى العالم الأمريكى المعاصر جيلفورد فى زيادة حركة قياس التفكير الابتكارى، فقد أثرت نظريته عن تنظيم العقل فى معظم بحوث الابتكار، وفيها يقترح تصنيف القدرات العقلية إلى فئتين كبيرتين هى قدرات التفكير، وقدرات الذاكرة، وتنقسم عوامل التفكير إلى ثلاثة أنواع هى: قدرات التفكير المعرفى (بمعنى الاكتشاف)، وقدرات التفكير الإنتاجى وقدرات التفكير التقويمى، ثم تنقسم قدرات التفكير الإنتاجى وقدرات التفكير التقويمى، ثم تنقسم قدرات التفكير الإنتاجى إلى

فئتين فرعيتين هما: قدرات التفكير الإنتاجى التقاربى، وقدرات التفكير الإنتاجى التباعدى، والإنتاج التقاربى (أو الاختزالى) يتضمن تضييق الاحتمالات عند انتاج إجابة واحدة محتملة للمشكلة، مثل: طويل إلى قصير مثل مرتفع إلى...؟ وبالطبع فإن أغلب اختبارات الذكاء العام تقيس هذا النمط من التفكير، بل أن معظم الاختبارات العقلية التى أشرنا إليها فى الفصول السابقة من هذا الكتاب تنتمى إلى فئة الإنتاج التقاربى، أما التفكير التباعدى فيتطلب إنتاج أكبر عدد ممكن من الإجابات . مثل : ما هى الاستخدامات التى يمكن أن تفكر فيها لقالب الطوب؟».

ويرى جيلفورد أن الإنتاج التباعدى عامل هام فى التفكير الابتكارى، واستخدم فى وضع اختبارات عديدة لقياس هذه القدرة فى مستويات مختلفة وفى محتويات متباينة، ومن رأيه أن بعض عوامل الذاكرة والتفكير المعرفى والتفكير التقويمى والتفكير الإنتاجى التقاربى ترتبط بالعمل الابتكارى، ويعتقد على وجه الخصوص أن الموهبة الابتكارية يمكن قياسها بالاختبارات التى تقيس قدرات: الأصالة والطلاقة والمرونة من مختلف الأنواع، وقدرات إعادة التحديد والتضمين والتفكير التقويمى أيضاً.

وقد أثرت نظرية جيلفورد فى ظهور بطاريتين رئيسيتين لقياس التفكير الابتكارى إحداهما أعدها جيلفورد وزملاؤه بجامعة كاليفورنيا، والأخرى أعدها تورانس بجامعة مينيسوتا.

اختبارات جيلفورد للتفكير الابتكارى:

تتألف بطارية اختبارات التفكير الابتكارى لجيلفورد على النحو الآتى:

١ – اختبارات جيلفورد- كريستنسن للطلاقة وتتألف من ٤ اختبارات :

(أ) طلاقة الأفكار (الإنتاج التباعدى لوحدات الرموز): يطلب فيه من المفحوص أن يكتب كلمات ذات خصائص معينة، كأن تتضمن حروفا معينة (ب مثلا).

(ب) طلاقة الأفكار (الإنتاج التباعدى لوحدات المعانى): يطلب فيه من المفحوص أن يذكر أسماء الأشياء التى تنتمى إلى فئة معينة (مثل السوائل التى تحترق).

(جـ) طلاقة التداعى (الإنتاج التباعدى للعلاقات بين المعانى)، يطلب فيه من المفحوص إعطاء كلمات تتشابه فى المعنى مع كلمة أخرى.

(د) الطلاقة التعبيرية (الإنتاج التباعدى لمنظومات الرموز): يطلب فيه من المفحوص تكملة جمل، كل منها يتألف من ٤ كلمات، بحيث تبدأ كل كلمة بحرف معين.

٢ – اختبار الاستخدامات البديلة (الإنتاج التباعدى لفئات المعانى): يطلب فيه من المفحوص أن يذكر الاستخدامات المحتملة لشيء معين غير استخدامه المألوف (قالب الطوب مثلا يستخدم فى البناء استخداما مألوفا).

٣ – اختبار المترتبات (الإنتاج التباعدى لوحدات المعانى، والإنتاج التباعدى لتحويلات المعانى). وفيه يطلب من المفحوص أن يذكر النتائج المختلفة المترتبة عن حادث افتراضى (مثل : ماذا يحدث لو أن الناس توقفوا عن النوم؟)، ويعطى للمفحوص فى هذا الاختبار درجتان: إحداهما للاستجابات الواضحة (وتقيس طلاقة الأفكار أو الإنتاج التباعدى لوحدات المعانى). والأخرى للاستجابات البعيدة (وتقيس الأصالة، والإنتاج التباعدى لتحويلات المعانى).

٤ – الأعمال المحتملة (الإنتاج التباعدى لتضمينات المعانى). ويطلب فيه من المفحوص أن يذكر الأعمال المحتملة التى يمكن أن يرمز لها بشعار ما (مصباح كهربائى مثلا).

٥ – عمل الأشياء (الإنتاج التباعدى لمنظومات الأشكال). ويطلب فيه من المفحوص أن يرسم أشياء محددة باستخدام مجموعة معطاة من الأشكال (كالدوائر أو المثلثات)، وعلى المفحوص ألا يستخدم إلا الأشكال المعطاة فقط.

٦ – الاسكتشات (الإنتاج التباعدى لوحدات الأشكال): ويشبه اختبار عمل الأشياء فيما عدا أن كل صفحة فى الاختبار تشتمل على الشكل نفسه (دوائر مثلا فقط) ويطلب من المفحوص أن يرسم أكبر عدد من الاسكتشات بإضافة تفاصيل إلى كل شكل من الأشكال المعطاة.

٧ – مشكلات عيدان الثقاب (الإنتاج التباعدى لتحويلات الأشكال): وفيه يطلب

من المفحوص أن يستبعد عدداً معينا من عيدان الثقاب (المرسومة على الورق فى صورة خطوط مستقيمة) بحيث ينتج عن ذلك عدداً من المربعات أو المثلثات.

٨ – اختبار الزخرفة (الإنتاج التباعدى لتضمينات الأشكال): وفيه يطلب من المفحوص أن يزخرف صوراً لأشياء مألوفة بأكبر عدد من الرسوم المختلفة.

ويلاحظ أن الاختبارات السبع الأولى من هذه القائمة تتطلب استجابات لفظية، بينما تستخدم الاختبارات الأربع الأخيرة محتوى من الأشكال أو الصور، وهذه الاختبارات الأحد عشر تمثل عينة صغيرة من مجموعات الاختبارات التى أنتجتها بحوث جيلفورد وزملائه وتلاميذه بجامعة جنوب كاليفورنيا لأكثر من عشرين عاما، وقد أعدت معايير مبدئية لهذه الاختبارات فى صورة مئينيات ومعايير جيمية C-Score (وهو معيار يلخص المستويات التائية إلى ١١ مستوى وهو فى ذلك يشبه المعيار التساعى ويختلف لعينات من الراشدين أو طلاب من سن ١٥ + أو كليهما، وتصلح هذه الاختبارات بوجه عام لمستوى المرحلة الثانوية وما بعدها، رغم أن جيلفورد يذكر أن بعض اختباراته تصلح لمستوى المرحلة الابتدائية (ابتداء من الصف الأول).

والمشكلة الجوهرية فى اختبارات التفكير الابتكارى هى مشكلة التصحيح فهى تتطلب من المصحح تدريبا كافيا على اتباع تعليمات التصحيح ودراسة الأمثلة التى تتضمنها كراسات التعليمات، ورغم ذلك فإن عملية تصحيح هذه الاختبارات تستغرق وقتا كبيراً.

وحسب جيلفورد ثبات هذه الاختبارات بطريقة التجزئة النصفية، وتتراوح معاملات الثبات بين ٠,٦٠ و، ٠,٨٠، أما الصدق فقد اعتمد اعتمادا كبيراً على صدق التكوين الفرضى كما يتمثل فى اتفاق الاختبارات مع نموذج العقل الذى يقترحه، ولازالت هذه الاختبارات فى حاجة إلى أنواع الصدق الأخرى وخاصة الصدق المرتبط بالمحكات.

وقد قام عبد السلام عبد الغفار (١٩) بإعداد بعض اختبارات جيلفورد باللغة العربية، فأعد اختبارين لقياس الطلاقة اللفظية (طلاقة الكلمات) هما اختبار الطلاقة اللفظية ١، ويتطلب من المفحوص إنتاج كلمات تبدأ بحرف معين، واختبار الطلاقة

اللفظية ٢ ويتطلب من المفحوص أن ينتج كلمات تنتهى بحرف معين، كما أعد ٣ اختبارات أخرى لقياس بعض عوامل التفكير الابتكارى وهى: اختبار الطلاقة الفكرية (لقياس عامل طلاقة الأفكار)، واختبار الاستعمالات (لقياس عامل المرونة التلقائية)، واختبار المترتبات (لقياس عامل الأصالة). وقد حسب ثبات هذه الاختبارات من أداء عينة عددها ١٢٠ تلميذا بالمدارس الثانوية العامة بطريقة إعادة الاختبار لاختبارى الطلاقة اللفظية فبلغ معاملات الثبات ٠,٦٢ و، ٠,٦٦ واستخدم فى حساب ثبات اختبارات الطلاقة الفكرية والاستعمالات والمترتبات طريقة التجزئة النصفية وبلغت معاملات الثبات ٠,٧٥، ٠,٦٩، ٠,٨٠، على التوالى، وفى تقدير الصدق حسب الباحث معاملات ارتباط الاختبارات بالتحصيل المدرسى، وتراوحت هذه المعاملات بين ٠,٢٧ (الاستعمالات) ٠,٥٨ (المترتبات)، ولم تتضمن كراسة التعليمات بيانات عن معايير الاختبارات.

كما قام سيد خير الله (١٣) بإعداد بعض اختبارات جيلفورد أيضًا وصدر له اختبار القدرة على التفكير الابتكارى يتكون من خمسة أقسام تقيس: الاستعمالات والمترتبات ولعبة الكلمات، ولا تتوافر لدينا أى معلومات سيكومترية عن هذا الاختبار.

اختبارات تورانس للتفكير الابتكارى:

أشرنا إلى أن اختبارات جيلفورد هى محصلة بحوثه العاملية فى ميدان التفكير الإنسانى وطبيعة الذكاء، أما اختبارات تورانس للتفكير الابتكارى التى نتناولها الآن فقد ظهرت لمواجهة الضرورات التربوية والعملية كجزء من برنامج طويل من البحوث تهتم بالخبرات التعليمية التى تساعد على تنمية الابتكار، وبعض اختبارات تورانس هى فى الواقع تعديل للأساليب التى استخدمها جيلفورد، بالإضافة إلى أن الدرجات المشتقة من البطاريات الكاملة تعتمد على العوامل التى توصلت إليها بحوث جيلفورد، وهى الطلاقة والمرونة والأصالة وإعطاء التفاصيل، وفى هذا الصدد يذكر تورانس أنه لم يسع إلى وضع اختبارات «نقية عامليا» أو «بسيطة التكوين»، وإنما إلى التوصل إلى اختبارات تتضمن مواقف لما تتطلبه العملية الابتكاريةفى الظروف الطبيعية والمعقدة، ولذلك فإن كل اختبار قد يصحح فى ضوء عاملين، أو ثلاثة عوامل، أو العوامل الأربع جميعا.

وتتألف اختبارات تورانس للتفكير الابتكارى من ١٠ اختبارات، مصنفة إلى بطاريتين: أحدهما لفظية والأخرى مصورة، وتسمى البطارية الأولى: التفكير الابتكارى بالكلمات، والثانية: التفكير الابتكارى بالصور، وحتى يمكن التخفيف من التهديد الذى تتضمنه كلمة «اختبار» استخدمت فى كل الأحوال كلمة «نشاط» كما تؤكد التعليمات على «اللعب» وتوصف هذه الاختبارات بأنها تصلح لمختلف المستويات ابتداء من مرحلة الحضانة حتى مستوى الدراسات العليا، بشرط أن تطبق فرديا وشفويا على المستويات الأدنى من الصف الرابع الابتدائى، وتوجد صورتان متكافئتان لكل بطارية من البطاريتين.

وفى «التفكير الابتكارى بالكلمات» الصورة (أ) و الصورة (ب) تتألف الأنشطة الثلاث الأولى (اسأل وخمن) من صورة غير عادية يستجيب لها المفحوص بكتابة (أ) جميع الأسئلة التى يحتاج إلى توجيهها لكى يعرف ما يحدث فى الصورة (ب) الأسباب المحتملة لما حدث فى الصورة (جـ) النتائج المترتبة على ما حدث فى الصورة، أما النشاط الرابع فيتعلق بطرق تحسين لعبة من لعب الأطفال بحيث تصبح مصدرا لمزيد من السرور والفرح لمن يلعب بها، ويتعلق النشاط الخامس بذكر الاستخدمات غير العادية لشىء مألوف: (علب الكرتون) فى الصورة (أ)، و(علب الصفيح) فى الصورة (ب)، ويتعلق النشاط السادس بالأسئلة غير المعتادة التى يمكن توجيهها عن الشىء المتضمن فى النشاط السادس بالأسئلة غير المعتادة التى يمكن توجيهها عن الشىء المتضمن فى النشاط الخامس، أما النشاط السابع فقد أعد على نسق اختبار المترتبات عند جيلفورد، ويتطلب من المفحوص أن يذكر كل ما يمكن أن يحدث إذا نشأ موقف غير ممكن الحدوث وتعطى البطارية درجات كلية فى العوامل الثلاث: الطلاقة والمرونة والأصالة.

ويتألف اختبار «التفكير الابتكارى بالصور»: الصورة (أ) والصورة (ب) من ثلاثة أنشطة:

(١) تكوين الصورة وفيه يعطى للمفحوص ورقة ملونة ذات شكل منحنى وعليه أن يضعها على صفحة بيضاء فى الموضع الذى يشاء ويستخدمها كبداية لرسم صورة غير عادية (والورقة الملونة المستخدمة فى الصورة (أ) زرقاء اللون، والمستخدمة فى

الصورة (ب) صفراء اللون فى الطبعة العربية.

ويحصل المفحوص فى البطارية المصورة على ٤ درجات هى: الطلاقة والمرونة والأصالة والتفاصيل.

وقد أعد هذه البطاريات باللغة العربية (بموافقة مؤلفها تورانس) عبدالله محمود سليمان وفؤاد أبو حطب (٢٠) وظهرت عام ١٩٧١ على النحو الآتى:

١ – التفكير الابتكارى بالكلمات: الصورة أ

٢ – التفكير الابتكارى بالكلمات: الصورة ب

٣ – التفكير الابتكارى بالصور: الصورة أ

٤ – التفكير الابتكارى بالصور: الصورة ب

٥ – كراسة تعليمات اختبارات تورنس للتفكير الابتكارى.

٦ – استمارة تصحيح (١) لاختبارات التفكير الابتكارى بالكلمات (الصورة «أ» أو الصورة «ب»).

٧ – استمارة تصحيح (٢) لاختبارات التفكير الابتكارى بالصور (الصور «أ» والصورة «ب»).

٨ – استمارة تقدير المعلمين للابتكار.

٩ – استمارة تقدير التلاميذ للابتكار.

وفى عام ١٩٧٧ نشر معدا الاختبار تقريرهما الأول (٢١) عن تقنين هذه البطارية، ويشمل على وجه الخصوص تقنين الصورة أ من اختبارات الأشكال أو الصور(*)، فطبق الاختبار على ٣١٣ تلميذا وتلميذة من مدرستين إعداديتين بالقاهرة تتراوح أعمارهم بين ١٢، ١٦ سنة، ولحساب الصدق استخدمت استمارة تقدير المعلمين للابتكار حيث طلب من مدرسى الفرق التى أجريت عليها الاختبارات أن يختار كل منهم ٥ تلاميذ يقعون فى الفئة العليا، ٥ تلاميذ يقعون فى الفئة الدنيا فى كل من أبعاد الابتكار المقيسة وهى: الطلاقة، المرونة، الأصالة، التفاصيل، واختير كل تلميذ يتفق عليه مدرسان على الأقل، وحسبت دلالة الفروق بين متوسطات مجموعات

(*) شكر المؤلفون كل من ساهم فى تصحيح الاختبارات التى طبقت فى البحث المشار إليه يومئذ السادة (الدكاترة فيما بعد) محمود السيد عمر، على السيد سليمان، سيد محمود الطواب، محمد متولى غنيمة، والأستاذ محمد السعيد عبدالحليم.

التلاميذ الأعلى والأدنى فى مختلف أبعاد الاختبار باستخدام مقياس (ت) فوجدت الفروق فى معظم الأحوال (باستثناء التفاصيل) دالة وفى الاتجاه الصحيح إلا أن مستوى الدلالة كان عاليا (أ. و) بالنسبة إلى بعد المرونة، ومنخفضاً نسبيا (٠,١٠) بالنسبة إلى بعدى الطلاقة والأصالة.

وقد استخدم فى حساب الثبات الطرق التالية:

١ – حساب معاملات الارتباط بين متغيرات الاختبار (الطلاقة، المرونة، الأصالة، التفاصيل، الدرجة الكلية) فوجدت المعاملات فى كل الأحوال عالية وذات دلالة إحصائية.

٢ – حساب معاملات الارتباط بين مصححين قاما بتصحيح ٣٨ اختبارا كل على حدة، وكانت معاملات ثبات التصحيح ٠,٩٩ للطلاقة، ٠,٩٩ للمرونة ٠,٩٩ للأصالة، ٠,٩٧ للتفاصيل، ٠,٩٩ للدرجة الكلية، ويرجع ارتفاع المعاملات على هذا النحو إلى التزام المصححين بقواعد مقننة فى التصحيح، كما تم تدريب المصححين على طريقة تورانس، وتأكدت بالفعل جدواها.

٣ – حساب الثبات بطريقة إعادة الاختبار (ن = ٣٨) وبلغت المعاملات ٠,٥٤ للطلاقة، ٠,٣٦ للمرونة، ٠,٥٣ للأصالة، ٠,٤٠ للتفاصيل، ٠,٥٠ للدرجة الكلية فى الاختبار.

وقد تم اعداد معايير الاختبار فى صورة درجات تائية، كما أعدت النماذج التجريبية لتصحيح الصورة (أ) من اختبارات الأشكال، وهى نماذج مشتقة من استجابات عينة التقنين المصرية ذاتها وليست منقولة عن كراسات التعليمات الأجنبية.

ولسوء الحظ لم تتوافر لمعدى الاختبار باللغة العربية فرصة أخرى للعكوف على الاختبارات الثلاثة الأخرى فى البطارية لتقنينها، كما لم تتح لهما الفرصة لإعادة تقنين اختبار الأشكال (الصورة أ) الذى مر على تقنينه الأول الآن أكثر من خمس وعشرون سنة، إلا أن مما يخفف من وقع الأمر أن الاختبارات الأربعة استخدمت فى بعض البحوث الامبريقية وفى عدد من رسائل الماجستير والدكتوراه فى الجامعات المصرية وعدد من جامعات العالم العربى وتؤكد نتائجها صلاحيتها للاستخدام ويرجو مؤلفو الكتاب امدادهم بالبيانات السيكومترية عن الاختبار كلما أمكن ذلك.

اختبار التفكير الابتكارى للأطفال:

أعدت هذا الاختبار فى أصله سيلفيا ريم S. Rimm عام ١٩٧٦ بهدف قياس التفكير الابتكارى عند الأطفال ويطلق عليه اسم بطارية الاختبارات الجماعية للكشف عن الموهوبين Group Inventory for Finding Talent (GIFT).

ويتكون من جزأين رئيسيين: أولهما ويصلح للأطفال من سن ٦ – ٩ سنوات، وثانيهما ويصلح للأطفال من سن ٩ – ١٢ سنة ويتضمن جزءا مصورا من اختبارات تورانس (الدوائر والمربعات) وجزءا لفظيا من اختبارات جيلفورد (الاستعمالات)، وبالإضافة إلى جزء من عبارات تتضمن صفات الطفل المبتكر.

وقد أعد الاختبار باللغة العربية سيد خير الله ومحمود عبد الحليم منسى (١٤)، وقاما بتقنينه على عينة من ٥٠٠ طفل من الذكور والإناث من تلاميذ المدارس الابتدائية بمدينة الاسكندرية، وحسب الثبات بطريقة إعادة الاختبار وتراوحت معاملاته لمختلف صفوف المرحلة الابتدائية بين ٠,٥١ ، ٠,٨٣ ، وفى حساب الصدق استخدام الباحثان تقديرات المعلمين للتفكير الابتكارى عند تلاميذهم كمحك، كما استخدم أيضاً اختبار القدرة على التفكير الابتكارى من إعداد سيد خير الله كمحك آخر.

اختبار التفكيرالابتكارى عند الأطفال باستخدام الحركات والأفعال:

أعد هذا الاختبار فى أصله الأجنبى تورانس Torrance وجبس Gibbs عام ١٩٧٧ ويعتمد فى جوهره على استخدام الحركات والأفعال ليكون مناسبا للأطفال بين ٣ – ٧ سنوات، ويتناول أربعة أنشطة على وجه الخصوص هى: كم طريقة؟ وتقدر تتحرك مثل:، هل فيه طريقة أخرى؟ الاستعمالات. وتقدر لهذه الأنشطة درجات للطلاقة والتخيل والأصالة.

وقد أعد الاختبار باللغة العربية محمد ثابت على الدين (٣٠)، وأعد له معايير تصحيح ملائمة للعينات المصرية، وطبقه على مجموعات من الأطفال المقيدين بدور الحضانة وبعض المدارس الابتدائية بالمنصورة عام ١٩٨١ ، وحسب له الثبات بطريقة إعادة الاختبار، وكانت المعاملات مرضية، وبذل الباحث محاولات جادة للعثور على محكات يحدد فى ضوئها صدق الاختبار، ولم يصلح استخدام الرسم كمحك، كما لم تصلح تقديرات المشرفين المتصلين بأطفال الحضانة، وكان أنسب المحكات هو تمايز العمر.

اختبار الابتكار فى مجال الموسيقى

أعدت هذا الاختبار د. آمال أحمد مختار صادق عام ١٩٨٤ كضرورة من ضرورات التعرف على المبتكرين فى مجال الموسيقى. فقد أشار جيلفورد فى مناقشته للإنتاج الابتكارى فى الفن (جيلفورد ١٩٥٧ : ١١٠ – ١١٨) إلى أن القدرات التى توصل اليها فى الميدان اللفظى – والتى تغلب على الاختبارات الراهنة للتفكير الابتكارى – ترتبط أكثر بالتأليف الأدبى، أما الابتكار فى الميادين الأخرى فيعتمد على قدرات أخرى تختلف تبعا لطبيعة المحتوى. ومعنى ذلك أنه فى حالة اختلاف طبيعة المحتوى للإختبار (لفظى أو رسوم) عن المحتوى المراد الابداع فيه (موسيقى) فإن هذا يؤدى إلى حصول الأفراد على درجات منخفضة (فؤاد أبو حطب – القدرات العقلية ١٩٩٦ :٢٤) ففى الابتكار اللفظى تكون الوحدة الاساسية للسلوك هى الكلمة، وفى الابتكار فى الفنون التشكيلية يكون الشكل هو وحدة السلوك، أما الموسيقى فتكون الوحدة الاساسية للسلوك الموسيقى هى النغمة (آمال صادق ١٩٦٨ : ٦١).

وبالطبع فإن النغمات باعتبارها وحدات للسلوك الموسيقى ففيها ترتبط معا لتظهر الجمل الموسيقية وهذا الارتباط قد يكون تتابعيا Successive وتتوافر فيه مسافات لحنية ويعد هذا أساسا للجمل التوافقية Harmony ومحتوى اختبار الابتكار الموسيقى يعتمد على الجانب اللحنى لانه أقرب إلى ذوق دارس الموسيقى فى بلادنا.

وصف الإختبار:

تتطلب مقاييس الإبتكار الموسيقى قدرة المفحوص على إكمال الجمل الموسيقية حيث قدم للمفحوص ثلاث جمل مبدئية Initial phrases فى مقامات وأزمنة مختلفة والزمن المحدد لكل جملة عشر دقائق تتكرر فيها الجملة نفسها عدة مرات، أى أن الزمن الكلى للإختبار ٣٠ دقيقة وتحسب درجة المفحوص على أساس استجابته للجمل الثلاث كالآتى:

١– طلاقة النغمات:

وهو العدد الكلى للنغمات التى تتألف منها استجابات المفحوص للجمل الموسيقية الثلاث بعد ان صنفت إلى فئات تكرارية تشمل النغمات الكوماتيكية Chromatic notes التى يشتمل عليها السلم الموسيقى.

٢- طلاقة الألحان: (الأفكار أو الجمل الموسيقية)

وهو العدد الكلى للإستجابات اللحنية التى يصدرها المفحوص للجمل الثلاث.

٣- المرونة الموسيقية:

حسبت المرونة الموسيقية على أساس عدد التحويلات التى يجريها المفحوص فى استجابته إلى مقامات Modesتختلف عن تلك التى صيغت فيها الجملة الأصلية Modulation وكذلك على مدى استخدام المفحوص للنغمات الكرماتيكية.

٤- الاصالة الموسيقية

حسبت درجة الأصالة على أساس مقياس تقدير خماسى مبتدأ من صفر إلى أربعة وقدرت درجة كل جملة أنتجها المفحوص بواسطة ثلاثة من الخبراء من المتخصصين فى مجال التأليف الموسيقى ثم حسب متوسط التقديرات واعتبر هذا المتوسط مؤشرا على أصالة الإستجابات.

صدق وثبات الاختبار:

طبق هذا الاختبار على ٥٦ طالبا وطالبة وهم جميع طلاب البكالوريوس فى كلية التربية الموسيقية فى العام الجامعى ١٩٨٤ واقتصر على هذه العينة حيث تبين ان مستويات الطلاب فى الفرق الدراسية الأدنى لا تعطى تباينا فى ادائهم وقد تم اخضاع نتائجهم للتحليل العاملى، حيث حسبت معاملات الارتباط بين المتغيرات الاربعة وتراوحت بين ٠,٣٤ ، ٠,٧٣ وأظهر التحليل العاملى عاملا واحدا اعلى تشبعاته فى طلاقة النغمات ٠,٨٢ ثم طلاقة الالحان ٠,٨٧ ثم الاصالة ٠,٦٥ ثم المرونة ٠,٣٨ وجميع هذه المعاملات دالة مما يؤكد الصدق العاملى للإختبار أم الثبات فتمثل فى تحديد اشتراكيات العوامل وكان اعلاها ٠,٧٢ لطلاقة النغمات . ٠,٧١ لطلاقة الالحان، ٠,٥٢ للأصالة. ٠,٢٦ للمرونة. وفى دراسة مستقلة للمقياس تم التحقق من استقلالية الابتكار الموسيقى عن الابتكار اللفظى كما تحدد بالأداء على اختبار جيلفورد للإبتكار إعداد عبد السلام عبد الغفار (١٩٦٥) وأكدت النتائج استقلالية نوعى الأداء (آمال صادق ١٩٨٥).

اختبار الشخصية المبتكرة

صمم دوجلاس هولمز (فى ٢٩) مقياس الشخصية المبتكرة Aquestionaire Measure of Creative Personality للكشف عن وجود سمات الشخصية المبتكرة لدى طلاب المرحلة الثانوية والجامعية. وقد اشار مؤلف المقياس الى ان نتائج الدراسات السابقة أكدت على أهمية دراسات الشخصية والتى اكدت على مفهوم الشخصية المبتكرة.

وقد أشار نسكا (١٩٨٥) إلى وجود انظمة فرعية متكاملة خاصة بالشخصية المبتكرة تتألف من نظام فرعى خاص بالبنية اللامعرفية، ونظام فرعى آخر يرتبط بالبنية المعرفية بما تشتمل عليه من قدرات عقلية، وعلى ذلك يمكن القول ان الابتكارية يمكن اعتبارها محصلة لمجموعة من العوامل المعرفية والمزاجية والدافعية.

ويتألف المقياس من ٧١ عبارة نصفها موجب والنصف الآخر سالب، وتندرج الدرجة على كل عبارة على مقياس خماسى، والدرجة التى يحصل عليها المفحوص هى المجموع الحسابى لاستجابات المفحوص على كل المفردات.

وقد قام مجدى عبد الكريم حبيب سنة ١٩٩٠ (٢٩) بتعريب المقياس وتقنينه على عينة من المرحلة الثانوية والمرحلة الجامعية قوامها (١٢٠٠ طالب وطالبة) نصفهم من الذكور والنصف الآخر من الإناث، ونصف العينة الكلية من المرحلة الثانوية والنصف الاخر فى المرحلة الجامعية.

وقام معد الاختبار من التأكد من صحة الترجمة،وبعد التطبيق ثم حساب الصدق التلازمى مستخدما قائمة السمات الابتكارية لجوردون، وقائمة الانشطة الابتكارية لتورانس، وأشار معد المقياس انه استخدم ايضا طريقة المقارنة الطرفية إلا أنه لم يحدد المحك المستخدم لتحديد المجموعات الطرفية .

أما الثبات فقد تم تحديده بإعادة التطبيق (ن = ١٠٠) من كل فئة فرعية (ذكور/ اناث، مرحلة ثانوية/ مرحلة جامعية) وتراوحت معاملات الثبات بين ٠,٤٨ ، ٠,٦٨ ، كما استخدم طريقة التجزئة النصفية، وتراوحت معاملات الارتباط بين ٠,٦٨ ، ٠,٨٩ وطريقة الفا كرونباك وتراوحت المعاملات بين ٠,٧١ ، ٠,٧٩ .

أما المعايير فقد تم تحويل الدرجات الخام إلى درجات تائية (درجة معيارية

معدلة متوسطها ٥٠ وانحرافها المعيارى ١٠ .

مقياس التفكير الأخلاقى للأطفال

أعدت هذا المقياس فوقية عبد الفتاح ٢٠٠١، وهو مشتق عن اختبار الأحكام الاخلاقية Moral Judgement test والذى اعده بياجيه فى اطار نظريته عن النمو الاخلاقى عند الاطفال، وطوره رونالد جونسون ١٩٦٢ . وقد خضع المقياس لدراسات متعددة قام بها كولبرج.

والمقياس يتكون من ٢٠ قصة تستثير اهتمامات الاطفال وعلى مستوى متكافىء فى مستوى صعوبتها وقد قام طلعت منصور غبريال وحليم بشاى بترجمته عام ١٩٨٢ .

ويتألف المقياس الحالى من خمس قصص مشتقة من المقياس الاصلى وتشمل مجالات الاحكام الاخلاقية الخمس كما حددها بياجيه، ويتم عرض المقياس المعدل مصورا لتيسير فهمه. والمقياس يصلح للأطفال من سن ٤-١٣ سنة، حيث تعرض القصة ويعقبها ثلاثة اسئلة يتم الاجابة عنها بطريقة الاختيار من متعدد (٣ بدائل) وذلك بوضع علامة(صح) أمام الاستجابة التى يختارها الطفل. والمفردات الخمس التى تقيسها المفردات هى

١- الجزاء فى مقابل العقاب التعويضى.

٢- المسئولية الجماعية.

٣- العدالة التلقائية

٤- قوة تأثير العقاب.

٥- الواقعية الخلقية.

ويتم تصحيح استجابات الاطفال بإعطاء درجة واحدة للاستجابة الناضجة لما تتضمنه القصة من مواقف تمثل التفكير الاخلاقى فى مقابل (صفر) للاستجابةغير الناضجة.

تقنين المقياس:

تم عرض المقياس على محكمين للتأكد من صحته، وكذلك على الاطفال فى

الفئات العمرية للتأكد من وضوح التعليمات وتحديد الزمن اللازم للاجابة سواء فى حالة التطبيق الفردى أو الجمعى. وقد تم تطبيق المقياس على ٢٠٠ طفلا (٥٠ طفل فى كل مجموعة عمرية فى الفئات العمرية (٤: ٦)، (٦: ٨) ، (٨: ١٠)، (١٠: ١٣) وعدد البنات فى كل مجموعة عمرية يكافىء عدد الذكور، وتم اختيار العينة من رياض الاطفال والمدارس الابتدائية والاعدادية بمحافظة بنى سويف وقد تم حساب الثبات بطريقة اعادة تطبيق الاختبار بفاصل ١٥ يوما وذلك فى العام الدراسى ١٩٩٩– ٢٠٠٠ ، وبلغ معامل الثبات ٠,٦٧ وهو معامل دال.

كما تم تحديد صدق المقياس باستخدام التحليل العاملى حيث تشبعت بالعامل الأول جميع مفردات المقياس، والعامل الثانى تشبعت جميع المفردات ماعدا المفردة الثالثة ولم يتضمن دليل المقياس أى اشارة عن وجود معايير.

اختبارات واطسون وجليزر للتفكير الناقد:

أعد هذا الاختبار جودون واطسون وادوارد جليزر ليزود المفحوص بمجموعة من المواقف والمشكلات التى تتطلب استخدام التفكير الناقد، ويتكون من ٥ اختبارات فرعية.

١ – الاستنتاج: ويقيس القدرة على التمييز بين الدرجات المختلفة من الصواب والخطأ، والتوصل إلى استنتاجات معينة على أساس البيانات المعطاة.

٢ – التعرف على الافتراضات: ويقيس القدرة على التعرف على الافتراضات المتضمنة فى القضايا المعطاة.

٣ – الاستنباط: ويقيس القدرة على التفكير الاستنباطى من مقدمات معينة.

٤ – التفسير: ويقيس القدرة على الحكم على الشواهد والأدلة، وللتمييز بين التعميمات التى تبررها الأدلة والتى لا تبررها.

٥ – تقويم الحجج: ويقيس القدرة على التمييز بين الحجج القوية والضعيفة بالنسبة لمسائل مطروحة أمام المفحوص.

وقد أعد هذا الاختبار فى صورته العربية جابر عبدالحميد جابر ويحيى حامد هندام (١٢).

وقد استخدم الاختبار فى عدد من البحوث ورسائل الماجستير والدكتوراه التى تناولت التفكير الناقد وتوافرت عنه بيانات سيكومترية كافية فيما يلى بعض نتائجها:

فقد قام أحمد محمد حسن (١أ) صالح بحساب صدق الإختبار بتحديد معامل الاتساق الداخلى (صدق المفردات) كما تحققت غادة تمساح (٢٣أ) من صدق المقياس باستخدام محك الإتساق الداخلى وذلك بتقدير معامل الارتباط بين كل مفردة والدرجة الكلية لكل بعد تنتمى له المفردة، وكذا الدرجة الكلية للمقياس وأبعاده الفرعية حيث تراوحت معاملات الارتباط بين ٠,٤٨ ، ٠,٧١ .

وقد قامت ايمان سعيد محمد مصطفى (١١أ) بدراستها فى الماجستير سنة ٢٠٠٥ بتطبيق المقياس على طلاب كلية التربية جامعة حلوان (ن = ٦٨) وقامت بتحديد صدق المقياس أيضا مستخدمة فى ذلك حساب معامل الارتباط بين اختبار واطسون وجليزر وكان المحك المستخدم هو اختبار التفكير الناقد لابراهيم وجيه المعدل وبلغ معامل الارتباط ٠,٤١ أما ثبات الاختبار فقد تم حسابه بإعادة التطبيق وكانت عينة التقنين قوامها ٧١ طالبا وبلغ معامل الارتباط ٠,٣٧ .

وقد ارجعت الباحثة انخفاض قيمة معامل الارتباط الى المفاهيم النظرية الخاصة بنظرية الحمل المعرفى Cognitive Load إذ احتوى اختبار واطسون وجليزر ٩٩ مفردة، مما شكل عبئا معرفيا على طلاب عينة التقنين استشعرته الباحثة أثناء التطبيق.

وفى دراسة نعمة عبد السلام محمد (٣٧ أ) فى دراستها للدكتوراه سنة ٢٠٠٨ والتى استخدمت ذات المقياس على طلاب كلية التربية بالسويس، حيث قامت بتحديد صدق المقياس أيضا من خلال حساب الصدق المرتبط بالمحك حيث بلغت قيمة معامل الارتباط بين اختبار تقويم الحجج واختبار ادراك التناقد الداخلى ٠,٢٩٩ وهذه القيمة بالرغم من انخفاضها دالة عند مستوى ٠,٠١ كما قامت ذات الباحثة بتحديد صدق المفردات باستخدام طريقة الاتساق الداخلى حيث تم حساب معامل الارتباط بين درجة المفحوصين على كل مفردة والدرجة الكلية للبعد الذى ينتمى له بعد تلقية الدرجة الكلية من درجة كل مفردة، وتم استبعاد المفردات التى لم يصل فيها قيمة معامل الارتباط مستوى الدلالة.

وتشير الباحثة إلى ضرورة تنقية الاختبار من المفردات التى لا تناسب الثقافة المصرية.

بطارية اختبارات الاستعدادات الحسية الحركية للمكفوفين:

أعد هذه البطارية فرج عبدالقادر (٢٢) لقياس الاستعدادات الحسية الحركية، للمكفوفين، وتتألف من الاختبارات الآتية:

١ – اختبار تمييز اللمس.
٢ – اختبار تمييز الأبعاد.
٣ – اختبار تمييز الوزن.
٤ – اختبار المهارة اليدوية.
٥ – اختبار مهارة الأصابع.
٦ – اختبارتآزر اليدين.
٧ – اختبار تصنيف الشكل ـ اللمس.

ولحساب الثبات طبقت الاختبارات على عينة من ٣٥ من المكفوفين وأعيد التطبيق عليهم بفاصل زمنى لا يقل عن شهر وكانت معاملات الارتباط دالة فيما عدا اختبارمهارة الأصابع، وحسب صدق التكوين الفرضى فى ضوء ما هو متوقع نظريا عن الفروق بين الجنسين، واستخدمت معايير العمر العقلى والدرجات المئينية والدرجات المعيارية.

مقياس التذكر المباشر:

من إعداد وفية الكمونى (٣٩) فى رسالتها للدكتوراه عام ١٩٧٦ ويتكون من اختبار للتعرف واختبار الاسترجاع بواسطة مثير مساعد، واختبار الاسترجاع الحر المباشر واختبار سعة الأرقام، وحددت الصدق كما تقول فى ضوء «التقدير الشخصى أوالاعتماد على آراء الحكام والخبراء المشهود بكفاءتهم»، وهى طريقة لايعتد بها كثيرا فى حساب الصدق. وقد تفيد فى المراحل الأولى لبناء الاختبار، إنما لا تعد طريقة كافية تصل بالاختبار إلى درجة من الصلاحية يعتمد عليها، كما أن الطريقة التى شرحتها الباحثة فى حساب الثبات (٣٩ : ٤ – ٧) غير واضحة، وعلى هذا لانوصى باستخدام هذا الاختبار بسبب ضعف معالمه السيكومترية.

اختبار الأشكال المتضمنة :

هذا الاختبار من إعداد وتكن Witkin وزملائه لقياس أحد الأساليب المعرفية التى شاعت فى السنوات الأخيرة وهو بعد الاعتماد على المجال فى مقابل الاستقلال عنه.

وقد نقل هذا الاختبار إلى العربية أنور الشرقاوى وسليمان الخضرى الشيخ (١١) وحسب ثباته بطريقة التجزئة النصفية من أداء عينة من طلاب جامعة الزقازيق (ن = ١٦٥) وبلغ معامل الثبات للبنين ٠,٧٦، وللبنات ٠,٧٨ كما استخدمت طريقة جتمان فى حساب الثبات وتكاد تكون القيم الإحصائية متطابقة فى الحالتين.

وأعدت للاختبار معايير فى صورة السباعى المعيارى، وقد أجريت على الاختبارات بحوث كثيرة تعد نتائجها شواهد على صدق التكوين الفرضى له.

مراجع الباب الرابع

١ – أحمد زكى صالح: اختبارالقدرات العقلية الأولية، مكتبة النهضة المصرية (ب. ت).

١ أ– أحمد محمد حسن صالح: الارتقاء فى المستوى الدراسى واثره على قدرات التفكير الناقد لدى الطلاب بالمرحلتين الثانوية والجامعية – مجلة كلية التربية – جامعة المنصورة – العدد ٢٥ الجزء الأول – مايو ١٩٩٤ .

٢– اسحق حنا بطرس: مدى التوافق بين اختبار كليات الهندسة والقدرات اللازمة للنجاح فيها، رسالة ماجستير، كلية التربية جامعة عين شمس، ١٩٧١ .

٣ – آمال أحمد مختار صادق: تقنين اختبارات ونج للذكاء الموسيقى على البيئة المصرية فى: بحوث فى تقنين الاختبارات النفسية (تحرير فؤاد أبو حطب) المجلد الثانى، مكتبة الأنجلو المصرية، ١٩٧٩، فى (٩) دراسات وبحوث فى سيكولوجية الموسيقى والتربية الموسيقية . الانجلو المصرية ٢٠٠٦ .

٤ – ـــــــــــ : مقاييس جديدة للقدرة الموسيقية، فى دراسات وبحوث فى سيكولوجية الموسيقى والتربية الموسيقية الطبعة الثانية – الانجلو المصرية ٢٠٠٦ .

٥ – ـــــــــــ: كراسة تعليمات اختبارات سيشور للقدرات الموسيقية، مكتبة الأنجلو المصرية. ١٩٧٢، ١٩٨٨، ٢٠٠٦، (فى ٩) .

٦ – ـــــــــــ : إعداد صورة مختصرة من بطارية سيشور، هى المرجع نفسه.

٧– ـــــــــــ: كراسة تعليمات اختبارات القدرات الموسيقية للأطفال «بنتلى»، مكتبة الأنجلو المصرية. ١٩٧٢، ١٩٨٨ ، ٢٠٠٦، (فى ٩).

٨– آمال أحمد مختار صادق: دراسة عاملية للإبتكار الموسيقى فى بحوث ودراسات فى سيكولوجية الموسيقى الانجلو المصرية ٢٠٠٦ .

٩– آمال أحمد مختار صادق: بحوث ودراسات فى سيكولوجية الموسيقة والتربية الموسيقية ، الانجلو المصرية (ط١) ١٩٩٤، (ط٢) ٢٠٠٦ .

١٠ – ـــــــــــ : إكرام مطر: القيمة التنبؤية لاختبارات القبول للتعليم الموسيقى فى مصر، المجلة الاجتماعية فى مصر، المجلد ١٣، العدد ٣، سبتمبر ١٩٧٦ .

١١ – أنور محمد الشرقاوى، سليمان الخضرى الشيخ: اختبار الأشكال المتضمنة. مكتبة لأنجلو المصرية، ١٩٨٥ .

١١ أ- ايمان سعيد محمد مصطفى: العلاقة بين التفكير الناقد وكل من الفهم اللغوى وبعض متغيرات الشخصية لدى طلاب الجامعة. رسالة ماجستير غير منشورة - كلية التربية جامعة حلوان ٢٠٠٥ .

١٢ - جابر عبد الحميد جابر، يحيى حامد هندام: كراسة تعليمات اختبارات التفكير الناقد، دار النهضة العربية، ١٩٧٠ .

١٣ - سيد محمد خير الله: اختبار القدرة على التفكير الابتكارى (ب. ت) .

١٤ - سيد محمد خير الله، محمود عبدالحليم منسى:اختبار التفكير الابتكارى للأطفال (ب. ت) .

١٥ - السيد محمد خيرى: علم النفس الصناعى: دار النهضة العربية ١٩٦٧ .

١٦ - السيد محمد خيرى: بطارية اختبارات استعدادات حرف المعادن ، مصلحة الكفاية الإنتاجية والتدريب المهنى (ب. ت) .

١٧- خالد يحيى مندوه : العلاقة بين كل من الذكاء الموسيقى ومفهوم الذات الموسيقى ومستوى التحصيل على آلة البيانو رسالة ماجستير غير منشورة - كلية التربية النوعية جامعةعين شمس ٢٠٠٣ .

١٨- ــــــــــــ: العلاقة بين كل من القدرات الموسيقية وبعض الاساليب المعرفية ومستوى الأداء على آلة البيانو - رسالة دكتوراه غير منشورة - كلية التربية النوعية - جامعة عين شمس ٢٠٠٦ .

١٩ - عبد السلام عبد الغفار: اختبارات القدرة على التفكير الابتكارى، مطبعة البيان العربية، ١٩٦٥ .

٢٠ - عبدالله سليمان، فؤاد أبو حطب: كراسة تعليمات اختبارات تورنس للتفكير الابتكارى، مكتبة الأنجلو المصرية، ١٩٧١ .

٢١ - عبدالله محمود سليمان، فؤاد أبو حطب: تقنين اختبارات تورانس للتفكير الابتكارى على البيئة المصرية، اختبارات الأشكال (الصورة أ) . فى بحوث تقنين الاختبارات النفسية، (تحرير فؤاد أبو حطب) ، المجلد الأول، مكتبة الأنجلو المصرية، ١٩٧٧ .

٢٢ - عبدالوهاب محمد كامل: اختبار هيبكى نبراسكا لقياس الاستعداد للتعلم، طنطا: المكتبة القومية الحديثة (ب، ت) .

٢٣ – على حسن بدارى، أنور رياض عبد الرحيم: الاختبارات المعرفية مرجعية العوامل (غير منشورة).

٢٣أ– غادة حسن تمساح: العلاقة بين التفكير الناقد والتفكير الابتكارى وبعض متغيرات البيئة الاسرية والمدرسين – رسالة ماجستير غير منشورة – كلة التربية بالاسماعيلية – جامعة قناة السويس ٢٠٠٦ .

٢٤ – فؤاد أبو حطب، القدرات العقلية، مكتبة الأنجلو المصرية (ط٤)، ١٩٨٣ . (ط٥) ١٩٩٦ .

٢٥ – ـــــــــ : التفضيل الفنى وسمات الشخصية، المجلة الاجتماعية القومية، يناير ١٩٧٣ .

٢٦ – ـــــــــ : آمال صادق: كراسة تعليمات اختبار مينيسوتا للوحات الأشكال ٢٠٠٧ (يطلب من معدى الإختبار لأغراض البحث العلمى فقط).

٢٧– فتحى الزيات: تقنين بطارية اختبارات فلاناجان على طلاب كلية الطب، رسالة ماجستير، كلية التربية جامعة المنصورة، ١٩٧٨ .

٢٨ – فرج عبد القادر طه: اختبارات الاستعدادات الحسية الحركية للمكفوفين، مطبعة دار التأليف، ١٩٧٤ .

٢٩– فوقية عبد الفتاح: مقياس التفكير الأخلاقى للأطفال – كراسة التعليمات الأنجلو المصرية ٢٠٠١ .

٢٩أ– مجدى عبد الكريم حبيب: اختبار الشخصية المبتكرة – كراسة التعليمات – دار النهضة المصرية ١٩٩٠ .

٣٠ – محمد ثابت على الدين: اختبارات تورانس للتفكير الابتكارى عند الأطفال باستخدام الحركات والأفعال، مكتبة الأنجلو المصرية، ١٩٨٥ .

٣١ – محمد عبد السلام أحمد، القدرة الكتابية واختباران لقياسها: لجنة التأليف والترجمة والنشر، ١٩٦٠ .

٣٢ – محمد عماد الدين إسماعيل، قياس الإبداع الفنى وعلاقته بالتربية الفنية، النهضة المصرية، ١٩٥٩ .

٣٣ – محمد عماد الدين إسماعيل، سيد عبد الحميد مرسى: كراسة تعليمات اختبارات المهن الكتابية، مكتبة النهضة المصرية (ب. ت).

٣٤ – محمود السيد أبو النيل: الحد الأدنى اللازم لأداء الاختبارات النفسية للسائقين، المركز القومى للبحوث الاجتماعية والجنائية، ١٩٧٥ .

٣٥ – محمود عبدالقادر: تقنين بطارية الاستعدادات العامة، المركز القومى للبحوث الاجتماعية والجنائية، ١٩٧١ .

٣٦– مرفت محمد عبد الله: اعادة تقنين اختبار القدرات العقلية الأولية– رسالة ماجستير غير منشورة – كلية التربية جامعة حلوان ٢٠٠٢ .

٣٧– مصطفى خليل الشرقاوى: دراسة لأنماط القدرات اللازمة للتكيف والنجاح فى التدريب على بعض المجموعات المهنية وتطبيقاتها فى مجال التربية والتوجيه المهنى، رسالة دكتوراه ، كلية التربية جامعة عين شمس، ١٩٧٣، أدلة بطارية اختبارات الاستعدادات الفارقة.

٣٨ – هدى برادة، فاروق صادق: بطارية اختبارات القدرات النفسية اللغوية وتقنينها على عينة مصرية فى كتاب: بحوث فى تقنين الاختبارات النفسية (تحرير فؤاد أبو حطب)، المجلد الثانى مكتبة الأنجلو المصرية، ١٩٧٩ .

٣٩ – وفية الكمونى:مقياس التذكر المباشر ومتغير المدى، مطبعة النهضة الجديدة ١٩٧٦.

40. anastasi, A. Psychological Testing. Macmillan, 1982.

41. Anastasi, A. Drake, J. An Empirical Comparison Of Certain Techniques For Estimating The Reliability Of Speeded Tests. Educ. Psycho. Measurement, 1954, 14, 529-540.

42. Anastasi, A And Urbina, S., Psychological Testing, Prentice Hall, 7th Edit. 1997 .

43. Balma, N. J., The Concept Of Synthetic Validity. Personal Psychology, 12, 395-396. 1959.

44. Bennett, G.k., Et Al. Test Of Mechanical Comprehension, Newyork, Psychol. Corporation, 1940-1954.

45. Crawford, J.e. & Crawford. D.m., Crawford Small Parts Dexterty 1981. In Anastasi; A., & Urbina, S. 1997.

46. Cronbach, L.j., Essentials Of Psychological Testing, Harper, 1984.

47. Flanagan, J. C. Flanagan Aptitude Classification Tests. Technical Report, Chicago, Sra, 1959.

48. Fleishman, E.a. Dimentional Analysis Of Movement Reactions, J. Exp. Psychol. 1958, 17, 522-532.

49. Freeman, F.s. Theory And Practice Of Psychological Testing. (3rd Ed.). Holt, Rinehart, Winston, 1962.

50. Guilford, J.p. , Crative Abilities In Arts. Rev., 64, 110-118, 1957.

51. Guilford, J.p. The Nature Of Human Intelligence, Mcgraw- Hill, 1967.

52. Likert, R,, Quasha, W.h. Revised Minnesota Paper Form- Board. Psychol. Corporation, 1941-1984.

53. Mellenbruch,.P.I. Mechanical Motivation Test. Psycho. Affiliates, 1957.

54. Melton, H.w. (Ed.) Apparatus Tests. Washington, D.c. Government Printing Offices, 1947.

55. Mourad S.a. & Moustafa M.m. The Development Of An Arabic Aptitude Test. A.u.c., Egypt.

56. Necka, E., On The Nature Of Creative Talent. In The 6th World Conference Of Gifted And Talented Children, Hamburg 1985 .

57. Nunnally, J.c., (Jr.) Introduction To Psychological Measurement. Mcgraw-hill, 1970.

58. Primof, E.s., Emperical Validation Of J- Coefficiant. Personal Psychology, 12, 413-418. 1959.

59. ---------- & Hgde, L.d., Job Element Analysis In: Goel S. Edit., The Job Analysis Handbook For Business, And Industry, And Government. (Vol 2 P.p. 807-824), New York Wiley 1988.

60. Richardson, B. Et Al. Sra Mechanical Aptitudes, Chicago; Science Research Associate, 1949-1950.

61- Sadek, A. A. M., A Factor Analytic Study Of Musical Abilities Of Egyptian Students Taking Music As A Special Subject. Ph. D. Thesis London University, 1968.

62. Sadek, A. A. M., The Nature Of Music Creativity. 6th World Confernce On Gifted And Talanted Children Hamburg, 1985.

63. Sadek, A. A. M., The Nature Of Flexibitiy Of Musical Closure. 14th International Research Commission Seminor, Nagoya, Japan, 1992.

64. Sadek, A. A. M., The Relationship Between Musical And Visual Embeded Figures. 15th International Rescarch Commission Seminar. Miami, Florida, 1994.

65. Super, D.e. (Ed.) The Use Of Multifactor Tests In Guidance, Washington, Amer. Pers And Guidance Assoc., 1959.

66. Super, D.e., Crites, J.o. Appraising Vocational Fitness By Means Of Psychological Tests. Harper, 1962.

67. Taylor, I, A., The Nature Of Creative Process. Hasting House, 1959 .

الباب الخامس
اختبارات التحصيل

الفصل الحادى عشر

بناء الاختبار التحصيلى

يقصد ببناء الاختبار التحصيلى achievement test الأداة التى تستخدم فى قياس المعرفة والفهم والمهارة فى مادة دراسية أوتدريبية معينة أو مجموعة من المواد.

وهذه الاختبارات هى أكثر الأدوات شيوعا فى التقويم التربوى، وهى فى بلادنا هى الوسيلة الوحيدة التى تستخدم فى توجيه المتعلمين وانتقائهم، كما أنها فى أغلب الأحوال ما يمكن للمعلم أن يصنعه فى ممارسته اليومية للتدريس، ولذلك فإنها تحتل منزلة خاصة بين أدوات التقويم النفسى يجعل من المهم إفراد هذا الفصل لها بحيث يمكن تناولها بشىء من التفصيل حتى يمكن للمعلم، والسلطات التربوية عامة، أن تستفيد من المبادئ التى سوف نتناولها هنا فى إعداد هذه الاختبارات.

تحديد الأهداف التعلمية:

أشرنا فى الفصل الأول من هذا الكتاب إلى العلاقة بين التقويم النفسى والأهداف التربوية educational objectives ، والواقع أن هذه الأهداف تتفاوت من العبارات الشاملة التى تدل على الأغراض العامة والقيم والفلسفات التربوية إلى العبارات المحددة التى تبين أساليب أداء نوعية أوخاصة، ويشار إلى النوع الثانى فى علم النفس التربوى المعاصر بالأهداف التعليمية instructional objectives، وإذا كانت الأهداف التربوية يعوزها الوضوح فى أغلب الأحوال، فإن الأهداف التعليمية أكثر تحديدا لأنها تتناول النواتج النهائية لعملية التعلم والتعليم فى صورة أداء يمكن ملاحظته، وبالتالى فهى على درجة أكبر من الوضوح.

ويرى كثير من علماء النفس التربويين المعاصرين (١٠) أنه من الواجب صياغة الأهداف التعليمية فى عبارات صريحة explicit تدل على الأداء النهائى terminal performance لثلاثة أسباب هى:

١ – يؤدى تحديد الأهداف التعليمية تحديدا سلوكيا بهذا الشكل إلى توجيه المعلم فى تخطيط عملية التعليم فيمكنه أن يحدد الخطوات التى يجب أن يسير فيها المتعلم للوصول إلى الهدف، كما يمكن أن ينهض بأعباء جميع الاستجابات التى يجب أن تصدر عن المتعلم للوصول إلى الاستجابات المهمة النهائية.

ومن المؤكد أن أساليب التدريس تعتمد على الأهداف التعليمية والمدخلات السلوكية، ومع ذلك فإن من نتائج تخطيط المعلم هذا وبحثه قد يكون تعديل الأهداف بتغيير مقدار الزمن مثلا أو المستوى المتوقع للاتقان أو الموضوعات التى يجب تناولها فى التدريس، وبالطبع لايمكن لهذه التغيرات أن تحدث ما لم تتم صياغة الأهداف بصورة صريحة، وفى الوقت نفسه لا يستطيع أن يوائم بين أساليبه فى التعليم والفروق فى المدخلات السلوكية ما لم يحدد بشكل دقيق، ومعنى ذلك أن أهم أسباب وجوب صياغة الأهداف بشكل صريح ما يهيؤه ذلك للمعلم من إرشاد فى تخطيط أساليبه فى التدريس.

٢ – يفيد تحديد الأهداف التعليمية فى عبارات صريحة فى تقويم الأداء ومن المعروف أن الإهتمام بالصياغة الدقيقة للأهداف التعليمية بدأه فى الأصل العلماء المهتمون ببناء الاختبارات وتقويم المناهج المدرسية، فقد اكتشفوا أن استخدام العبارات الغامضة فى التعبير عن الأهداف يجعل عملية بناء الاختبار ومفرداته عملية صعبة إن لم تكن مستحيلة.

٣ – تفيد ترجمة الأهداف اتعليمية فى صورة أداء نهائى فى توجيه انتباه المتعلم وجهوده.

وقد أثبت ميجر ومكان (١٩) أن التعلم يمكن أن يحدث حين لا يفعل المعلم أكثر من إعطاء المتعلمين قائمة بالأهداف التعليمية الواضحة يقوم المتعلمون بالاهتداء بها فى تعلمهم الذاتى.

وصف العمل التربوى:

يتطلب تحديد الأهداف التعليمية وصفا دقيقا للعمل التربوى كما يتطلب تحليلا تفصيليا لهذا العمل، وطريقتا وصف العمل task description وتحليل العمل task analysis نشأتا أول الأمر فى ميدان علم النفس الصناعى ثم انتقلتا فى السنوات الأخيرة إلى ميدان علم النفس التربوى.

ومن أهم أساليب وصف العمل التربوى طريقة ميجر Mager وطريقة ميللر Miller، وتتطلب طريقة ميجر (٧، ١١) فى وصف العمل ـ أو بعبارة أخرى تحديد الأهداف التعليمية ـ خطوات ثلاث هى:

١ – تحديد الأداء النهائى الذى يسعى التعليم أو التدريس إلى إنتاجه.
٢ – وصف الظروف والأحوال المهمة التى نتوقع حدوث السلوك فيها.
٣ – وصف مستوى الجودة الذى يجب أن يصل إليه أداء المتعلم.

فمثلا قد يكون الهدف التعليمى الذى نسعى إلى تحقيقه هو أن يستطيع المتعلم تحديد العوامل التى أدت إلى حدوث الحرب العالمية الأولى، وهذا هدف مترجم فى أداء يمكن ملاحظته ملاحظة صريحة وعامة فى صورة أداء نهائى يسعى التعليم إلى إنتاجه (الخطوة الأولى)، ثم يمكن أن نعطى للمتعلم قائمة بالعوامل التى تؤدى إلى الأحداث التاريخية المهمة، وتكون هذه القائمة وصفًا لما يسميه ميجر العوامل المهمة التى يحدث فى ظلها السلوك النهائى (الخطوة الثانية)، وأخيرا تتحدد المستويات المطلوبة (الخطوة الثالثة)، والواقع أن تحديد المستويات هذا ليس من الأمور القاطعة ويتراوح بين الاتقان الكامل وبين أقل مستويات الاتقان.

وتشير طريقة ميللر (١٠) فى وصف العمل إلى وجود ثلاث خطوات أيضاً هى:

١ – ظروف الاستثارة stimulus condition أو ما يسميه indication.
٢ – التنشيط Activation أو الاستجابة التى يجب أن تصدر عن المتعلم.
٣ – التغذية الراجعة Feedback أو تحديد مستوى ملاءمة الاستجابة.

والواقع أن هذه الخطوات لا تختلف كثيرا عن تلك التى أشار إليها ميجر، فظروف الاستثارة هى ما يشير إليه ميجر بالعوامل أو الظروف أو الأحوال المهمة، والتنشيط عند ميللر يقابل ما يسميه ميجر بالأداء النهائى، ويتشابه مفهوم التغذية الراجعة عند الأول من مفهوم مستوى الأداء عند الثانى.

تحليل العمل التربوى:

بعد ما يقوم المعلم بوصف الأعمال التى يتوقع من المتعلمين أداءها يكون عليه تحليل هذه الأعمال وتصنيفها، والغرض الرئيسى من تحليل العمل هو مساعدة المعلم على تحديد أنسب ظروف التعلم، والتى تختلف من نوع من أنواع التعلم لآخر، فتعلم تكوين المفاهيم أو حل المشكلات أو الابتكار يتطلب ظروفا تختلف عن اكتساب المهارات الحركية أو حفظ قائمة من الكلمات. وهذا يعنى بالطبع أن مسئولية المعلم أكبر من مجرد وصف العمل التربوى أو صياغة الأهداف فى عبارات سلوكية

صريحة، فهى تشمل أيضًا تصنيف هذه الأساليب السلوكية المتضمنة فى كل هدف من هذه الأهداف ثم تحديد أنسب شروط التعلم لكل منها.

ويوجد فى الوقت الحاضر تصنيفان أساسيان للسلوك لهما أهميتهما التربوية هما تصنيف جانيه Gagné وتصنيف بلوم Bloom . ويشمل تصنيف جانيه (١٤) ثمانية أنواع من التعلم تتدرج من البسيط إلى المركب هى:

١ – التعلم الإشارى Signal learning أو ما يسمى فى سيكولوجية التعلم بالاشتراط الكلاسيكى وفيه يتم تعلم استجابة شرطية لمثير أو إشارة معينة.

٢ – تعلم المثير ـ الاستجابة أو ما يسمى بالاشتراط الإجرائى.

٣ – التسلسل chaining أو ما يسمى اكتساب المهارات الحركية (غير اللفظية).

٤ – الارتباط اللغوى verbal assoeiation.

٥ – التمييز المتعدد Multiple discrimination.

٦ – تعلم المفاهيم Concept learning.

٧ – تعلم المبادئ Principle learning.

٨ – حل المشكلة Problem solving أو التفكير.

أما التصنيف الذى وضعه بلوم Bloom ومساعدوه للميدان العقلى المعرفى (٩) فيشتمل على فئتين كبيرتين هما المعرفة Knowledge والقدرات والمهارات العقلية بحيث نحصل فى النهاية على ستة فئات للأهداف التربوية هى:

١ – المعرفة Knowledge.

٢ – الفهم Comprehension.

٣ – التطبيق Application.

٤ – التحليل Analysis.

٥ – التركيب Synthesis.

٦ – التقويم Evaluation.

وصف المدخلات السلوكية:

تتطلب عملية التعليم وصفًا للمدخلات السلوكية لدى للمتعلمين حتى يمكن للمعلم أن يخطط لتنفيذ عملية التدريس، بل وحتى يمكن للسلطات التربوية أن تخطط للعملية

التربوية عامة، ويتطلب هذا الوصف تحديدا دقيقا للوضع الراهن لمعارف المتعلم ومعلوماته ومهاراته وطرق تفكيره وقدراته واستعدادته وميوله ودوافعه وخصائص شخصيته، قبل أن نسعى إلى إكسابه النماذج السلوكية الجديدة من خلال التعلم والتدريس، وهذه البيانات يجب استخدامها لأغراض الوصف وليس لأغراض التفسير، لأن هذه المفاهيم تصف فى الواقع المستوى السلوكى للفرد فى لحظة معينة، ولا تستطيع الاختبارات والمقاييس التى نستخدمها فى قياس هذه الظواهر أن تفسر لنا أداء الفرد فيه فلا يمكن مثلا أن نعزو ضعف الأداء فى اختبار للذكاء إلى «ضعف الذكاء» لأن ذلك يصبح نوعا من تحصيل الحاصل ولايزيد فهمنا للسلوك الإنسانى.

والواقع أن وصف المدخلات السلوكية ضرورى لأى برنامج تربوى فعملية التعليم يجب أن تبدأ مما لدى المتعلمين ويجب أن تستمر إلى حيث يستطيعون من أداء نهائى تحت شروط معينة ولمستوى معين، كما أن تحديد أهدافنا يمكننا من أن نقرر ما سوف نعلمه وما نفترض أنه لدى المتعلمين بالفعل ويمكن للتفاعل بين الأهداف والمدخلات السلوكية مع وسائل التدريس وأساليبه، فنحن لانبدأ تدريسنا لتحقيق هدف معين بمواد تعليمية جديدة تماما وإنما قد نستدعى الاستجابات التى سبق تعلمها وخاصة تلك التى تتناسب مع موقف التعلم الجديد.

ثم إن لكل نوع من التعلم مدخلاته السلوكية فتنمية سلوك حل المشكلة تتطلب معرفة بمدخلات سلوكية تختلف عما يتطلبه تنمية إحدى المهارات الحركية (٦).

العلاقة بين الأهداف التعليمية وأساليب التقويم التربوى:

يلخص جرونلاند (١٥) العلاقة بين الأهداف التربوية والأهداف التعليمية وأساليب التقويم التربوى على النحو الآتى:

الأهداف التربوية العامة «الأهداف التى توجه التدريس توجيها يوميا»

↓

نواتج التعلم الخاصة أو النوعية «الأهداف التعليمية»

«سلوك المتعلمين وأداؤهم الظاهر الذى يعد دليلا عن تحقيق هذه الأهداف»

↓

أساليب التقويم

«الأساليب التى يحصل بها على عينة من سلوك المتعلمين كما توصف فى نواتج التعليم النوعية».

ويتطلب هذا ربط أساليب التقويم مباشرة بنواتج التعليم التى نقومها وبهذا نتيقن من أننا نقوم المتعلم نحو الأهداف التى أخذناها كإطار لعملية التدريس. وعملية ربط أساليب التقويم بنواتج التعلم النوعية هى فى أساسها عملية تحليل منطقى وحكم، هذه العملية يمكن تيسيرها باستخدام خطة ما من خطط التقويم المنظمة، هذه الخطة يجب أن تشمل على الأقل قائمة بنواتج التعلم المرغوب فيها والأساليب التى تستخدم فى تقويم التقدم نحوها، وإليك خطة على سبيل المثال فى ميدان علم الأحياء للصف الأول الثانوى تشمل الأهداف التعليمية ونواتج التعلم النوعية وأساليب التقويم الملائمة.

الأهداف	التقويم
١ – يعرف المتعلم المصطلحات الشائعة المستخدمة فى علم الأحياء حيث:	
١ – ١ يعرف المصطلحات الشائعة فى علم أساس المعنى	– اختبار موضوعى – الإجابة القصيرة.
١–٢ يميز بين المصطلحات الشائعة على أساس المعنى	– اختبار موضوعى.
١–٣– يتعرف على معانى المصطلحات الشائعة حين تستخدم فى نص معين.	
٢ – يظهر المتعلم القدرة على التفكير الناقد حينما:	

الاهداف	التقويم
٢ –١ يميز بين الحقائق والاراء	– اختبار موضوعى
٢ –٢ يشتق استنتاجات صحيحة من المعطيات	– اختبار مقال من نوع الإجابة القصيرة.
٢ –٣ يتعرف على الافتراضات وراء الاستنتاجات	– اختبار موضوعى.
٢ –٤ يتعرف على نواحى القصور فى المعطيات	– اختبار موضوعى
٣ – يؤدى المتعلم الأعمال الأساسية للتشريح بمهارة حينما:	
٣ – ١ يضع العينة فى الوضع الصحيح للتشريح	– قائمة ملاحظة أو مقياس تقدير
٣ –٢ يقطع بمهارة دون افساد العينة التى يدرسها	– قائمة ملاحظة
٣ –٣ يقسم أجزاء البنية فى العينة دون إفسادها	– قائمة ملاحظة أو مقياس تقدير.
٣ –٤ ينتهى من التشريح فى الوقت المحدد	– قائمة ملاحظة أو مقياس تقدير
٤ – يظهر المتعلم القدرة على تحديد مواضع المعلومات البيولوجية حينما:	
٤ –١ يستخدم كتالوج المكتبة فى تحديد المراجع	– تقرير بحث أو قائمة ملاحظة
٤ –٢ يعرف المصادر المشتركة للمعلومات البيولوجية	– اختبار موضوعى.
٤ –٣ يستخدم جداول المحتوى والفهرس عند البحث عن المعلومات فى الكتب	– قائمة ملاحظة.
٤ –٤ يحدد مدى ملاءمة المعلومات لمشكلة محددة	– تقرير بحث أو قائمة ملاحظة.
٥ – يكون لدى المتعلم اتجاه علمى نحو الظواهر البيولوجية حينما:	
٥ –١ يؤجل الحكم حتى تتوافر جميع الحقائق أو معظمها.	– اختبار موضوعى.
٥ –٢ يبحث عن علاقات السبب والأثر فى البيانات البيولوجية	– تقرير قصصى ، – اختبار موضوعى
٥ –٣ يظهر رغبة فى التفاسير الجديدة للبيانات البيولوجية	– تقرير قصصى – اختبار موضوعى
٥ –٤ يفسر البيانات البيولوجية متحررا من التحيز	– تقرير قصصى – اختبار موضوعى
٥ –٥ يظهر ثقة فى البيانات البيولوجية التى يحصل عليها بالأساليب العلمية	– تقرير قصصى، – اختبار موضوعى

من هذا الجدول نجد ما يأتى:

١ – نواتج التعلم النوعية كما تصاغ فى صورة سلوك للمتعلمين عديدة ومتنوعة

بحيث لاتوجد أداة تقويم واحدة تعطى دليلا مباشرا على تحقيق هذه الأهداف التربوية، ولذلك نحتاج إلى جانب الاختبارات الموضوعية إلى قوائم الملاحظة والتقارير القصصية واختبارات المقال.

٢ – تظهر أهمية الصياغة الواضحة للأهداف ونواتج التعلم فى انتقاء أداة التقويم، وفى الواقع حيثما تصاغ نواتج التعلم فى ضوء سلوك المتعلم فإنها لا تدل فقط على ماذا نقوم ولكنها توحى أيضاً بكيف نقوم.

مثال: (١ – ١) «يحدد المصطلحات الشائعة» يوضح لنا أسلوب التقويم الذى يجب أن يستخدمه، فهى تدل على أن المتعلم يجب أن يعطى التعاريف بنفسه وعلى ذلك فإن اختبار المقال من نوع الإجابة القصيرة حيث يعطى للمتعلمين عدة مصطلحات ويطلب منهم تعريفها هو الأسلوب الأكثر ملاءمة لتقويم هذا الهدف، أما الاختبار الموضوعى مثلا من نوع الاختيار من متعدد حيث يطلب من المتعلمين أن يتعرفوا فقط على التعريف لايتناسب مع تقويم هذا الناتج التعليمى فى صيغته هذه، وإلا فإن الصيغة تصبح «يتعرف المتعلم على معنى المصطلحات الشائعة»، وفى هذه الحالة يمكن استخدام الاختبار الموضوعى، ويتوقف هذا على غرضنا من التقويم، فإذا كنا نريد أن نتأكد من أن المتعلمين يعرفون المصطلحات عن طريق تحديدهم وتعريفهم لها فإن الإجراء الوحيد الملائم هو أن يطلب من المتعلمين أن يعرفوا مصطلحاتهم، والقدرة على التعرف الصحيح لا يمكن اعتبارها دليلا على قدرة المتعلمين على إعطاء التعريف الصحيح ما لم يكن هناك دليل من البحث العلمى يؤكد أن نوعى السلوك مرتبطان ارتباطا عاليا. وحيث أنه لايوجد مثل هذا الدليل بعد فإن الإجراءت الأحوط هو اختيار أسلوب التقويم الذى يرتبط مباشرة بنتائج التعلم النوعى (الهدف التعليمى).

تحليل محتوى المادة الدراسية:

أشرنا إلى أن الاختبار التحصيلى يرتبط بمادة دراسية معينة تم تدريسها بالفعل، ومعنى ذلك أنه لايوجد ما يبرر إعداد اختبارات من هذا القبيل تتناول مواد لم يتم تدريسها أو لم تكن محددة كمحتوى لعملية التعلم، ولذلك لابد فى إعداد الاختبار التحصيلى من تمثيل كل ما حظى باهتمامنا أثناء عملية التدريس ، ومن الطرق

الملائمة فى تحديد الأصل الإحصائى المناسب الذى نشتق منه عينة السلوك التى تسمى الاختبار التحصيلى أن نعد قائمة بالموضوعات التى تتضمنها هذه المادة بأكبر قدر من التفصيل وفى أبسط صورة ممكنة ، ويمكن للباحث فى هذا الصدد أن يستفيد بمنهج البحث المعروف باسم تحليل المحتوى، Content analysis (٥) .

تحديد الوزن النسبى للأهداف التعليمية ومحتوى المادة الدراسية :

بعد تحليل المحتوى يقوم الباحث بتحديد الوزن النسبى لكل موضوع (أو وحدة) من موضوعاته، ومن المحكات المستخدمة فى تحديد وزن كل موضوع من موضوعات المحتوى هو الزمن المخصص لتدريس كل منها، إلا أنه فى بعض الأحيان لايكون الوقت المستغرق فى تدريس موضوع معين دلالة على أهميته، ولذلك فإن الأوزان النسبية للموضوعات لابد أن تعكس اهتمام المعلم به. وبالمثل لابد من أن يحدد المعلم الأوزان النسبية للأهداف التعليمية لمادته، فإذا كان يؤكد فى تدريسه وصول المتعلمين إلى التعميمات وتطبيقها مثلا فيجب أن يعطى هذا الهدف وزنا أكبر، وبعبارة موجزة لابد أن تعطى لموضوعات المحتوى وللأهداف التعليمية أوزان نسبية تعكس الاهتمام الذى تحظى به فى عملية التدريس.

وفى تحديد هذه الأوزان قد يعتمد المعلم على خبرته فى تدريس المادة وتحليله لمحتواها فيضع قائمة بأهم الموضوعات ثم يصنفها حسب الأهداف التى يرى أن يقيسها الاختبار . إلا أنه فى كثير من الأحيان لا يكتفى الباحث بالاعتماد على خبرته الشخصية وحدها ويلجأ إلى بعض الخبراء لاستشارتهم فى تحديد هذه الأوزان النسبية .

إعداد جدول المواصفات:

على أساس المعلومات التى يتم تجميعها يتم تحديد مواصفات الاختبار test specifications والتى تتضمن الموضوعات التى يجب أن يشملها الاختبار ونواتج التعلم التى يجب اختبارها (أى الأهداف التعليمية، والأهمية النسبية (أو الوزن النسبى) للموضوعات والأهداف، وعلى هذاالأساس يتم تحديد عدد الأسئلة التى يتم إعدادها مرتبطة بكل هدف فى كل موضوع، والطريقة الأكثر سهولة فى تحديد مثل هذه المواصفات هى إعداد جدول المواصفات table of specifications كما تناولناه فى الفصل الخامس.

تحديد نوع الأسئلة:

أشهر أنواع الأسئلة التى تستخدم فى ميدان الاختبارات التحصيلية أسئلة المقال والأسئلة الموضوعية، وأسئلة حل المشكلات (أو الأسئلة التفسيرية)، ونعرض فيما يلى عرضا موجزا لأنواع الأسئلة الشائعة فى التقويم التربوى الحديث.

أسئلة الإختيار من متعدد:

وتعتبر أكثر الأسئلة الموضوعية شيوعا ، وتقيس بكفاءة شديدة النواتج البسيطة للتعلم، بل قد تصلح لقياس بعض النواتج التعليمية المعقدة، ويتكون السؤال من مشكلة (قد تصاغ فى صورة سؤال مباشر أو عبارة ناقصة) تسمى الجذر Stem ،وقائمة من الحلول المقترحة (قد تشتمل كلمات أو أعداد أو عبارات) وتسمى البدائل الاختبارية al- ternatives. ويطلب من المفحوص قراءة جذر السؤال وقائمة البدائل وانتقاء البديل الصحيح أو الأفضل، والبديل الصحيح فى كل سؤال يسمى ببساطة الإجابة answer بينما تسمى البدائل الاختبارية المحولات (أوالمشتتات) distractors، وهى تسمى كذلك من واقع وظيفتها المقصودة، أى تحويل هؤلاء الذين فى شك من الإجابة الصحيحة.

ويستخدم هذا النوع من الأسئلة فى قياس كثير من الأهداف التعليمية وخاصة نواتج التعلم فى ميدان التذكر واكتساب المعلومات، ومن ذلك معرفة المصطلحات ومعرفة الحقائق النوعية، ومعرفة المبادئ والقوانين،ومعرفة الطرق والإجراءات، كما تستخدم فى قياس الفهم مثل القدرة على تطبيق الحقائق والمبادئ، والقدرة على تفسير علاقات السبب والأثر، والقدرة على تبرير الطرق والإجراءات.

ويتميز هذا النوع بمرونته الشديدة، وتحرره، من نقائص الأنواع الأخرى من الأسئلة، ومن مشكلة أسلوب الاستجابة response set بل أن صياغة البدائل الاختيارية قد يؤدى استخدامها فى أغراض التشخيص، ومع ذلك فإن لأسئلة الاختيار من متعدد مشكلاتها وأهمها أنها تقتصر على المستوى اللفظى للأداء (وهى فى هذا تشترك مع غيرها من أسئلة الورقة والقلم)، كما أنها لا تصلح لقياس بعض مهارات حل المشكلة فى الرياضيات والعلوم، والقدرة على تنظيم الأفكار (وهى فى هذا تشترك مع غيرها من أسئلة التعرف على الإجابة)، ثم إن لأسئلة الاختيار من متعدد نقيصة خاصة بها لا تشترك فيها مع غيرها، وهى صعوبة إيجاد عدد كاف من المحولات

المحبذة وغير الصحيحة فى وقت واحد، إلا أن هذه الصعوبة تتناقص مع زيادة الخبرة فى بناء مثل هذه الأسئلة.

أسئلة التفسير:

يتطلب هذا النوع من الأسئلة ما هو أكثر من الحفظ واكتساب المعلومات الجزئية، أى أنه يستخدم فى قياس نواتج التعلم المرتبط بالعمليات العقلية العليا كالفهم ومهارات التفكير وقدرات حل المشكلة، ويتكون السؤال التفسيرى interpretative من سلسلة من الأسئلة الموضوعية تعتمد على مجموعة مشتركة من البيانات الأولية (المعطيات). وقد تكون هذه المعطيات فى صورة مواد مكتوبة أوجداول أو رسوم أو أشكال أوخرائط أو صور. وقد تتخذ الأسئلة المرتبطة صورا مختلفة ولكنها فى الغالب تأخذ صورة الاختيار من متعدد.

أسئلة المزاوجة:

تتكون أسئلة المزاوجة matching التقليدية من عمودين متوازيين يحتوى كل منهما على مجموعة من العبارات أوالرموز أو الكلمات، وتسمى المفردات فى العمود الذى ننشد له المزاوجة بالمقدمات premises، وتسمى المفردات التى نختار منها الاستجابات responses.

وعادة ما تكون أسس مزاوجة الاستجابات للمقدمات واضحة فى ذاتها، وفى بعض الأحيان تتطلب توضيحاً فى التعليمات، ويستخدم هذا النوع فى قياس الحقائق التى تستند إلى التداعى البسيط مثل الشخصيات والإنجازات، التواريخ والأحداث التاريخية، المصطلحات والتعاريف، القواعد والأمثلة، الرموز والمفاهيم، المؤلفون ومؤلفاتهم، الكلمات الأجنبية ومقابلاتها بالعربية ، الآلات واستخداماتها، النبات وتصنيفاته، الحيوان وتصنيفه، الأشياء وأسماؤها، الأعضاء ووظائفها، وتتميز بأنها تقيس بدقة وسرعة مقدارا كبيرا من الحقائق المرتبطة فى زمن قصير نسبياً، بالإضافة إلى أنها سهلة الصنع. ومع ذلك فإن لهذا النوع مشكلاته وأهمها أنه يقتصر على قياس الحقائق التى تعتمد على التذكر والتداعى البسيط، وكذلك يصعب على الباحث إيجاد مادة متجانسة لها أهميتها فى المحتوى الدراسى يمكن أن تصاغ منها قائمة المقدمات.

ويشترط فى بناء هذا النوع من الأسئلة أيضا أن يزيد عدد المفردات التى يتم الاختيار من بينها فى العمود الذي يتضمن الاستجابات (ويفضل أن يكون ثلث عدد

المفردات فى عمود المقدمات) حتى لا يكون تساوى عدد المفردات فى كلا العمودين يؤدين إلى التخمين وايضا يكون السؤال الاخير محلولا عند تساوى العدد.

أسئلة الإجابة القصيرة (الاستدعاء) :

يتطلب هذا النوع من الأسئلة الذى يعرف فى بعض المؤلفات باسم أسئلة التكملة - أن ينتج المفحوص استجابته - وليس مجرد التعرف عليها كما هو الحال فى الأنواع السابقة، ويمكن القول أن هذا النوع قد يتطلب إجابة قصيرة إذا عرضت المشكلة فى صورة سؤال مباشر، أو تكملة إذا عرضت فى صورة عبارة ناقصة، وتوجد أنواع أخرى من هذا النوع منها : أسئلة إعداد القوائم Listing والتى تسمى أحيانا اختبارات المقال ذات التعليمات المحددة (مثل اذكر ٣ أسباب لقيام الحرب العالمية الأولى) وأسئلة التماثل أو التناسب analogy، وأسئلة المشكلات ، وأسئلة التعيين identification (كأن يطلب من المتعلم تعيين أجزاء رسم نبات أو حيوان) . ويستخدم هذا النوع من قياس النواتج البسيطة نسبيا للتعلم مثل معرفة المصطلحات، ومعرفة الحقائق النوعية، ومعرفة المبادئ، ومعرفة الطرق والإجراءات ، والتفسيرات البسيطة، والقدرة على حل المسائل فى الرياضيات والعلوم، وهذا النوع ربما بسبب بساطة نتائج التعلم التى تقيسها (الاستدعاء فى أغلب الأحوال) ، بالإضافة إلى تحرره من أثر التخمين السائد فى اختبارات التعرف بوجه عام، إلا أن أكثر مشكلاته شيوعا هى صعوبة تصحيحه.

أسئلة الصواب والخطأ (أو أسئلة البديلين alternative response) :

وتتطلب اختيار إجابة واحدة من إجابتين كالحكم على العبارة بالصواب أو الخطأ، أو الإجابة على السؤال بنعم أو لا ، أو الحكم على العبارة بأنها تدل على رأى أو حقيقة، وتقدير عبارة بالموافقة والمعارضة، ويستخدم هذا النوع فى قياس نتائج التعلم التمييزى البسيط، ومع ذلك فيمكن أن تفيد فى قياس الفهم (كالتعرف على علاقات السبب والأثر) ، وقد يكون أكثر مشكلات هذا النوع من الأسئلة هو خطورة تأثرها البالغ بالتخمين ولذلك لابد فى جميع أنواع الاسئلة التى يتم فيها الاختيار من بدائل تصحيح الدرجة التى يحصل عليها المتعلم من أثر التخمين، وهذه المعادلة هي :

$$د = ص - \frac{خ}{ن-١}$$

حيث يدل الرمز د على الدرجة المصححة من أثر التخمين

الرمز ص على عدد الإجابات الصحيحة فى الاختبار

الرمز خ على عدد الاجابات الخاطئة فى الاختبار

الرمز ن على عدد البدائل الاختيارية فى الاختبار. فمثلا فى أسئلة الصواب والخطأ تكون (ن = ٢).

أسئلة الترتيب:

وفيها يقوم المفحوص بإعادة ترتيب خطوات أو أحداث أو إجراءات أو تواريخ فى تسلسل طبيعى أو منطقى.

أسئلة المقال:

رغم الاستخدام الواسع للأسئلة الموضوعية التى أشرنا إليها لا زالت توجد نتائج مهمة للتعلم لا يصلح لها إلا سؤال ، المقال Essay Type ، ومنها القدرة على عرض وتنظيم وتكامل الأفكار، والقدرة على التعبير الكتابى، والقدرة على إعطاء التفسيرات والتطبيقات للمعلومات لا مجرد استدعاؤها أو التعرف عليها.

ومن أهم خصائص أسئلة المقال حرية الاستجابة، وفى هذا جدواها كمقياس للتحصيل المعقد، وفيه أيضاً تكمن صعوبات التصحيح التى تجعل منها أداة أقل كفاية وكفاءة فى قياس الحقائق والمعلومات، ومعنى ذلك أن أسئلة المقال يجب أن تستخدم حيث لا تصلح الاختبارات الموضوعية.

الاختبارات العملية*:

تستخدم هذه الاختبارات كوسائل موضوعية لتقدير الكفاءة التى يؤدى بها العمل (كالكتابة على الآلة الكاتبة) أو قيادة السيارات، وتنقسم هذه الاختبارات إلى ٣ أنواع:

(أ) اختبارات التعرف: وتتطلب من المفحوص التعرف على الخصائص الأساسية للأداء (كأن تعزف مقطوعة موسيقية ويطلب من المفحوص تحديد الأخطاء أوالنقائص فى الأداء أو التعبير) أو تحديد الأجزاء التى تتألف منها احدى الآلات أو

* سبق الإشارة الى هذا النوع من الاختبارات فى الفصل السابق تفصيلا.

اختيار الالة ، أو الجهاز المناسب لعمل معين، أو تحديد العينات أو تصنيف الأشياء، أو انتقاء عمل أدبى أو فنى ممتاز.

(ب) الاختبارات التى تتضمن مواقف تشبه المواقف الطبيعية: فهى تهدف إلى قياس الأنشطة الأساسية فى العمل، وتسمى أحيانا اختبارات النماذج الصغيرة mini-ature.

(جـ) اختبار عينة العمل: وهو عبارة عن محاولة «مضبوطة» أو «مقننة» فى الظروف الواقعية للعمل، وتنقسم هذه الاختبارات إلى نوعين أساسيين: أولهما الاختبارات التى يسهل فيها التمييز بين الصواب والخطأ فى الأداء وبالتالى يمكن تصحيحها بسهولة مثل التصويب، والأداء العضلى فى التربية البدنية، والتجميع الميكانيكى، والكتابة على الآلة الكاتبة، والنوع الثانى هو الإختبارات التى تعتمد على حكم المراقبين والفاحصين لتقويم الأداء وإعطاء درجة أو رتبة، ومن أمثلة ذلك محاولة سامى إبراهيم على (٤) «تقنين اختبارات العزف على البيانو، ويتطلب هذا النوع استخدام مقاييس التقدير أو قوائم الملاحظة.

الاختبارات الشفوية*:

الاختبار الشفوى هو مزيج من اختبار المقال والاختبار العملى، وله فائدته فى دراسة العمليات العقلية التى يستخدمها المفحوص فى الإجابة على أسئلة معينة، ولذلك فهو أداة نافعة فى تشخيص الصعوبات، بل إنه فى بعض الأحوال هو الأسلوب الأوحد فى تقويم المفحوص كما هو الحال فى قياس قدرات صغار المتعلمين (فى الصفوف الأولى من المرحلة الابتدائية)، وفى قياس بعض نتائج التعلم اللغوى (كالقراءة الجهرية).

والإختبار الشفوى له تاريخ طويل فى تاريخ التربية، فحتى قرن واحد من الزمان كان الاختبار الشفوى هو الصورة المسيطرة على التقويم التربوى، وقد حدث التحول فى الوجهة Paradigm shift مع ظهور فنيات التقويم الحديثة التى حلت تدريجيا محل الإمتحان الشفوى، إلا أن له بعض الاستخدامات حتى وقتنا الحاضر لعل من

---

* عن فؤاد أبو حطب (بتصرف): الامتحان الشفوى: ماله وما عليه – ورقة القيت فى الندوة التى عقدتها جامعة أسيوط – وحدة تقويم الأداء الجامعى ١٧–١٨ ابريل ٢٠٠٠ فى المجلة المصرية للدراسات النفسية العدد التذكارى٣١ ابريل ٢٠٠١ .

أهمها مناقشة الرسائل الجامعية، كما يستخدم أحيانا فى امتحان بعض التخصصات لمزاولة مهنة معينة مثل التدريس حيث يصبح للإمتحان الشفوى قيمته فى التخصصات التى تتطلب التفاعل مع الآخرين، على إفتراض أن هذا التعامل يعد نوع من المهارات العملية والتى تتطلب فنيات خاصة فى تقويمها شأنها فى ذلك شأن أى مهارات عملية.

والاختبار الشفوى يمكن تعريفه على أنه علاقة مباشرة بين شخصين فاحص (ممتحن) ومفحوص (ممتحن) من نوع الوجه للوجه face to face وتكون مهمة الفاحص إما قبول أو رفض اجابة الطالب أو اللجوء إلى اسئلة التعمق (مثلما الحال فى المقابلة الشخصية) كما أن على الفاحص أن يحكم على نوعية إجابة المفحوص.

وموقف الاختبار الشفوى قد يصبح أكثر تعقيدا عندما يتألف موقف التقويم التربوى من أكثر من فاحص واحد (حرصا على الموضوعية) أو حين يكون موقف الإمتحان الشفوى من عدة مفحوصين. وعلى ذلك يصبح موقف الإمتحان الشفوى موقفا مركبا ويتألف من مستويات عديدة لكل منها وظيفته وأغراضه. فقد يكون مجرد إستجواب (سؤال وجواب)، أو مقابلة، أو اختبار أداء، أو يكون هذا كله. وبالطبع فإن الامتحان الشفوى مهما كانت صورته هو اتصال مباشر وتواصل إيجابى وتفاعل حى بين فاحص ومفحوص ويبدو أنه بهذه الصورة أكثر شخصية أو ربما أكثر انسانية من الامتحان التحريرى بشرط ان يتحرر من تهديد الذات للمفحوص بل ويصبح خبرة سارة وممتعة بين الفاحص والمفحوص.

مجالات استخدام الاختبار الشفوى:

تتحدد المجالات الأساسية لاستخدام الاختبار الشفوى فيما يلى:

١- مهارات التحدث والاستماع فى اللغة وهما أصل المهارات اللغوية جميعا، فالقراءة والكتابة مهارات أكثر حداثة فى تاريخ الإنسانية.

٢- اسئلة المعلم أثناء التدريس لطلابه، وأثناء المحاضرة أو خلال موقف المناقشة حيث يكون السؤال الشفوى وظيفته التعليمية.

٣- اسئلة المتابعة للإجابة التحريرية حين يطلب من الطالب التعليق على اجابته

أو توسيعها أو تبريرها، وخاصة فى حالات الشك فى الغش فى الإمتحان التحريرى أو عدم الفهم أو التعرف على عمليات الإجابة Answer Process حين يكون الناتج Product مضللا فى الحكم.

وبالطبع فإن للإختبار الشفوى حتى ولو أحسن توظيفه فى أغراضه المخصصة له بعض الحدود أهمها:

١- إضاعة الوقت وإستهلاكه بسبب طبيعته الفردية وخاصة إذا خصص للإمتحان الشفوى الوقت الواجب له (عادة لايقل عن نصف ساعة) حتى يمكن الحصول على عينة ملائمة من أداء الطالب.

٢- العلاقة الشخصية والمباشرة بين الفاحص والمفحوص قد تفتح الباب للتحيز أو اثر الهالة والانطباع الشخصى والذاتية على حساب تقويم عينات السلوك المطلوبة مما يؤثر على دقة التقويم.

٣- قد تؤثر المواجهة المباشرة بين الفاحص والمفحوص فى العمليات المعرفية لدى المفحوص والتى ربما تكون عادية فى الظروف المعتادة. وقد يتعارض ذلك مع بعض ما هو شائع من أن الإختبار الشفوى يختبر قدرة الطالب على التفكير فى الحال أو تحت الضغط والكبت. فمن حقائق علم النفس المعرفى الحديث أن الإنسان يؤدى أفضل أداء حينما لا يكون تحت ضغط التهديد أو الكبت، والإبداع أكثر العمليات المعرفية حاجة إلى ذلك.

٤- صعوبة الحصول على أحكام دقيقة إلا إذا توافرت شروط خاصة

ويقترح فؤاد أبو حطب فى ذات المصدر (٢٠٠٠) بعض الوصايا لزيادة فعالية الامتحان الشفوى فى أغراضه الخاصة وهى:

١- تجنب إستخدام الإختبار الشفوى إذا كان الإمتحان التحريرى يمكن أن يؤدى المهمة على نحو أكثر فعالية (مثل قياس النواتج المعرفيةCognitive products).

٢- تحديد أغراض استخدام الإختبار الشفوى والأساس الذى سوف يستخدم فى الحكم على أداء الطلاب. (تحديد السلوك موضع التقويم، إعداد بطاقات الملاحظة أو مقاييس، تحديد أنماط الاسئلة التى سوف تطرح على المفحوصين).

٣- استخدام فاحصين على الأقل بشرط أن يكونا من أهل الخبرة والثقة معا، ولهما الفة بموضوع التقويم والتدريب على استخدام الاختبار الشفوى وخاصة فى أغراض الامتحان والسمات التى تقوم، وأنواع الاسئلة، ومعنى درجات مقياس التقدير.

٤- تهيئة الفرصة للطالب لأن يؤدى الامتحان الشفوى فى جلستين مستقلتين على الأقل- كلما كان ممكنا - مع فاحصين مختلفين فى كل حالة.

٥- تدريب الطلاب على الأداء فى الإختبار الشفوى (معرفة أغراض الإختبار وطبيعته وأساس الحكم والتقدير فيه، والصيغة العامة للاسئلة التى تطرح، والإجراءات التى تستخدم. مع تجنب الاسئلة الخادعة أو غير المرتبطة أو تلك التى يستحيل حلها لمعرفة ما إذا كان الطالب يدرك ذلك) ولابد أن يكون ذلك ضمن تهيئة الطلاب للإمتحان.

٦- تنظيم البيانات لمراجعة صدق وثبات البيانات والأحكام، فالصدق هنا يبنى على أساس الأحكام التى يميزها الفاحصون، وحساب مؤشرات الموضوعية للتغلب على إختلاف الأحكام وتعادلها.

٧- ضرورة التنبيه إلى مزالق الإختبارات الشفوية، وطرق تحسينها لتؤدى أغراضها حتى لا تصبح الجهود المبذولة ضائعة فى سراب الصعوبة وضباب الذاتية وعدم الثبات والدقة.

تحديد عدد الأسئلة:

يتحدد عدد الأسئلة التى يتكون منها الاختبار بمقدار الزمن الميسر له. وكثير من الاختبارات التحصيلية يتحدد زمنه بحوالى ٥٠ دقيقة (زمن الحصة المدرسية العادية) وبعض جداول الامتحانات تحدد فترات أطول تصل إلى ٣ ساعات ، وعلى وجه العموم كلما طال الزمن وزاد عدد الأسئلة كان الاختبار أكثر ثباتا، ومن النادر من الوجهة العملية أن نعد اختبارا ليستغرق أكثر من ثلاث ساعات الا فى حالات قليلة مثل اختبارات الفنون التشكيلية والتى يتطلب من الطالب أن ينتج عملا كاملا مثلما يحدث فى كلية الفنون الجميلة.

ومن المفيد كما أشرنا أن نعتبر مجموعة الأسئلة التى يتكون منها الاختبار عينة

من أصل افتراضى لكل الأسئلة الممكنة التى يمكن أن تستخدم فى الاختبار، فمدرس الصف الخامس الابتدائى قد يحصل على ١٠٠ كلمة لاختبار المفردات اللغوية وذلك بانتقاء الكلمة الخامسة من قائمة تحتوى على ٥٠٠ كلمة درست أثناء العام الدراسى، وفى هذه الحالة تصبح الـ ٥٠٠ كلمة هى الأصل الذى تؤخذ منه عينة الاختبار، وهو أصل حقيقى ومحدد المعالم، إلا أنه فى معظم الاختبارات لا نجد مثل هذا التحديد الحاسم للأصل الإحصائى العام للأسئلة، فمثلا لايوجد حد أقصى لعدد المشكلات والمسائل المختلفة التى يمكن إعدادها فى اختبارات الجبر.

وبالطبع كلما اتسع مدى الموضوع أو القدرات التى يهدف الاختبار إلى قياسها زاد عدد الأصل الإحصائى، وحجم هذا الأصل يمثل الحد الأقصى لحجم عينة الاختبار، ومعنى ذلك أن أسئلة الاختبار لايمكن أن تكون أكثر عددا من الأصل التى تشتق منه، إلا أن حجم هذا الأصل لا يصنع حدا أدنى لحجم العينة، فالأصل الذى يتألف من ١٠٠٠ سؤال يمكن أن تكون العينة المشتقة منه ١٠ أسئلة أو ٥٠ سؤالا أو ٢٠٠ سؤال وكلما زاد حجم الأصل قل تجانسه، أى أصبح يحتوى ميادين مختلفة وشبه مستقلة من المعرفة والقدرة، وللوصول إلى نتائج دقيقة يجب أن تكون عينة الاختبار المشتقة من أصل غير متجانس أكبر نسبيا من تلك التى تشتق من أصل متجانس، ومع ذلك فلا توجد قاعدة عامة تحدد العلاقة من حجم الأصل والعينة.

ويوجد فى الوقت الحاضر اتجاه يؤكد استخدام عدد قليل وكاف من الأسئلة بحيث يعطى معظم المفحوصين وقتا يحاولون فيه جميع الأسئلة بمعدلاتهم الخاصة العادية ولشيوع هذا الاتجاه ثلاثة أسباب (١٣)هى:

١ – إن سرعة الاستجابة ليست هدفا أوليا للتعلم فى معظم المقررات الدراسية وبالتالى لا تعد مؤشرا صادقا للتحصيل لأنه فى كثير من ميادين التحصيل لا ترتبط السرعة والدقة ارتباطا عاليا، وبالطبع فقد تكون اختبارات السرعة أكثرملاءمة فى بعض الحالات، إلا أن هذه المواقف هى الاستثناء وليس القاعدة.

٢ – إن قلق الاختبار قد يكون على درجة كبيرة من الحدة ـ حتى فى الاختبارات غير الموقوتة ـ حيث يتزايد التأكيد على العمل بسرعة.

٣ – إن الاستخدام الكفء للاختبار التحصيلى ـ الذى يصنعه المعلم ـ يتطلب أن معظم المتعلمين يستجيبون له كله حتى تستخدم نتائجه فى أغراض التشخيص.

نماذج من التقويم فى المدرسة الإبتدائية

١- تقويم التحصيل فى القراءة فى المدرسة الابتدائية(*) :

قد تكون القدرة على القراءة أهم القدرات التى تؤكدها المدرسة الابتدائية، وفى الصفوف الأولى من المرحلة الابتدائية يحتل تعليم القراءة كمهارة أساسية الجزء الأكبر من العمل المدرسى، ومع انتقال المتعلم إلى الصفوف العليا تتفرع دراسته إلى مجالات متنوعة ويقل الوقت المخصص لتعليم القراءة، ومع ذلك فإن لها أهميتها طوال المرحلة الابتدائية، وبعدما يتقن المتعلم بعض المهارات الأساسية فى القراءة، وقد تتعدل تعديلات طفيفة أهداف تدريس القراءة بحيث تؤكد التغيرات الناتجة فى المهارات التى نقيسها، ومن ناحية أخرى فإنه لازال من الأهداف الأساسية للتعلم فهم معانى الكلمات وفهم الجمل والفقرات الطويلة، وهى جميعا لها أهميتها فى جميع المستويات، ومعنى هذا أن أى اقتراحات تتعلق بتقويم أهداف القراءة فى أحد صفوف المرحلة ينطبق على كل المرحلة.

وفيما يلى مثال لأهداف نوعية لتعليم القراءة فى الصف الأول الابتدائى:

الغرض العام	الهدف المرحلى (فى نصف العام)	الأهداف النوعية
أولا: أن يستطيع الفرد المتعلم قراءة اللغة القومية قراءة فعالة.	أ. أن يستطيع المتعلم قراءة القسم الأول من كتاب القراءة بسرعة وفهم.	١ – أن يستطيع المتعلم أن ينطق بكلمات مثل مدرسة، الحديقة، ورد، حينما يراها مكتوبة (عدد كلمات القسم الأول من القراءة ٢٥٠).
		٢ – أن يستطيع المتعلم أن يربط كلمات مثل ولد، طيارة، بيت وبين الصور التى تعبر عنها.
		٣ – أن يستطيع المتعلم أن يجيب على أسئلة تدور حول مواد جديدة يقرأها من مختارات من كتاب القراءة.
	ب. أن يستطيع المتعلم القراءة الجهرية بسهولة مع بعض التعبير.	٤ – أن يستطيع المتعلم أن يغير من نبرات الصوت للتمييز بين الجملة الخبرية والجملة الاستفهامية.

(*) الأمثلة التى سوف نعرض لها مأخوذة بكثير من التعديل عن لندفال (١٧).

ومن الواضح أنه لو كان الواجب هو تقويم هذه الأهداف النوعية جميعا فيجب أن تستخدم أساليب مختلفة، فالهدفان ١ ، ٤ يمكن تقويمها ببعض صور الأداء الشفوى، بينما الهدفان ٢ ، ٣ يمكن تقويمهما بهذه الطريقة أو باستخدام بعض المواد المكتوبة، وبسبب الفوائد الكثيرة الخاصة باستخدام الاختبارات الموضوعية فى المدرسة الابتدائية فإن على المعلم أن يختار لقياس تحصيل الهدف النوعى رقم ٢ طريقة المزاوجة بين صور الأشياء والكلمات التى تدل عليها، وفى هذا المستوى يجب أن يعطى لجميع المتعلمين تعليمات شفوية تقرر أنهم يجب عليهم أن يرسموا خطا يصل بين الكلمة والصورة التى تدل على معنى الكلمة، ومثل هذا السؤال يعطى قياسًا مباشرا لقدرة الطفل على ربط كل كلمة بالصورة الملائمة، والعلاقة البسيطة بين الهدف وهذا الاختبار تدل على فائدة صياغة الأهداف فى عبارات تدل على سلوك يمكن ملاحظته.

وبالنسبة لتقويم الهدف ٣ فإن المعلم يطلب من المتعلمين الانتقال إلى مواد جديدة وقراءتها قراءة صامتة، وبعد هذه القراءة توجه إليهم أسئلة يطلب منهم كتابة إجابات عليها وتتعلق هذه الأسئلة بمحتوى الفقرة، وهذا الاختبار قد يأخذ صورة أن يطلب من المتعلمين أن يكتبوا جملة فى الإجابة على السؤال - وهو صورة مبسطة لاختبار المقال- إلا أن المتعلمين فى هذا المستوى لانتوقع لهم أن تكون عندهم سهولة كبيرة فى الكتابة، ولذلك فإن بعض الاختبارات الموضوعية ينتج قياسا أكثر صدقا لفهم القراءة عندهم مثل أسئلة الاختيار من متعدد مثل:

فى يوم العيد ذهبنا إلى: الحقل

المدرسة

الحديقة

ومرة أخرى يجب أن تكون تعليمات هذه الأسئلة شفوية وتطلب من المتعلم أن يضع خطا أو يرسم دائرة حول الإجابة الصحيحة للسؤال.

وحيث أن القراءة هى مهارة اتصال، فإن جميع أهداف التعليم فى هذا الموضوع تهتم أساسا بالفهم، وخاصة فهم معانى الكلمات، فمن المتوقع من المتعلم أن يشرح معنى كلمة ، أو يختار مرادفا لها، أو يربط بينها وبين الصورة أوالشىء الذى تصفه، كما تربط الأهداف الأخرى بينهم الجمل أو الفقرات.

ومع تقدم المتعلم فى صفوف المرحلة الابتدائية فإن التأكيد على التحصيل فى القراءة يتزايد اهتمامه بأنماط مختلفة وجوانب مختلفة من الفهم، فتتطلب اختبارات الفهم من المتعلم أكثر من مجرد استدعاء أجزاء من معلومات الحقائق والتفاصيل.

٢- تقويم التحصيل فى الحساب:

تتزايد أهمية المهارات الحسابية والفهم الحسابى فى الحضارة المعاصرة تزايدا مستمرا، ومن المهم، للمتعلمين فى المرحلة الابتدائية أن يتلقوا تعليما جيدا فى المهارات الأساسية فى الحساب وفى عناصر التفكير الكمى، وللمساعدة فى تحقيق هذا الهدف من المهم للمعلم أن يقوم بتقويم مستمر لتقدم كل متعلم، وهذا مهم على وجه الخصوص فى الحساب حيث أن كل مهارة لاحقة تعتمد على المهارة السابقة، وحيث يفشل المتعلم إذا فشل فى إتقان أى من المعلومات الأساسية مثل:

الغرض العام	الهدف المرحلى	الهدف النوعى
أولا: أن يستطيع المتعلم حل مشكلات العد والحساب.	أ. أن يستطيع التلميذ أن يعد حتى رقم ١٠٠.	١ - أن يستطيع التلميذ أن يعد ويسجل العدد الصحيح لأشياء أوحروف توضحها الصور فى حدود ١٠ وحدات. ٢ - أن يستطيع الطفل العد إلى رقم ١٠٠ بزيادة واحد. ٣ - أن يستطيع التلميذ الإجابة على مسائل الجمع لأزواج من الأعداد التى تتكون من رقم واحد بحيث يكون مجموعها أقل من ١٠.
	ب. أن يعرف التلميذ حقائق الجمع الأساسية.	٤ - أن يستطيع التلميذ شرح معنى مصطلحات مثل جمع وطرح ومجموع وأكثر وأقل.
	جـ. أن يستطيع التلميذ فهم المصطلحات الحسابية الخاصة.	٥ - أن يستطيع التلميذ أن يختار اختيارا صحيحا للأشياء التى تتحدد بمصطلحات مثل الأول والثانى والثالث... إلخ.

وجميع الأهداف النوعية الخمسة يمكن تقويمها إما شفويا أو تحريريا، فالهدف يمكن قياسه باستخدام أسئلة بسيطة يعطى فيها للطفل أعدادا مختلفة من الأشياء عليه أن يعدها.

تعليمات (شفوية): اكتب على الخط الموضوع أسفل كل صورة العدد الذى يدل على عدد الحروف فى كل منها.

د د ص ص ح ح ه ه ه

د ص ص ص ح ح ه ه

ص ص ح ح ه ه ه

ويمكن تقويم الهدف الثانى بسؤال المتعلم أن يعد جهريا أو أن يكتب الأعداد من ١ إلى ١٠٠.

والهدف الثالث يمكن قياسه بكفاءة أكثر بمجرد سؤال المتعلم أن يكتب إجابته على عدد من المسائل التى تتضمن الجمع البسيط:

٢	٧	٥	١	٦
٤ +	٢ +	٤ +	٨ +	٣ +

والهدف الرابع يمكن تقويمه بسؤال المتعلم مباشرة أن يشرح هذه المصطلحات شفويا أوتحريريا، ويوجد تعديل طفيف لهذا الهدف بحيث يتطلب من المتعلم أن يتعرف على صور أشياء أو عمليات يمكن أن توصف لفظيا كأن يعطى له رسم لعدد من الكرات متفاوتة الحجم ويطلب منه وضع خط تحت الكرة الأصغر.

والهدف الخامس يمكن تقويمه باستخدام أسئلة مماثلة، كأن يعطى له رسم لعدد من الكرات المتساوية ويطلب منه وضع خط تحت الكرة الرابعة مثلا.

إن أنواع الأسئلة التى ذكرناها حتى الآن تتعلق أساسًا بقياس معلومات المتعلم وتذكره لحقائق خاصة أو مصطلحات نوعية، إلا أن المعلم قد يهتم بأنماط أخرى من القدرات وخاصة تلك التى ذكرناها فى تصنيف الأهداف التربوية (لبلوم) وتحت عناوين مثل الفهم والتطبيق.

مثال لترجمة الأعداد والمفاهيم إحداها إلى الآخر:

١ – ما الذى يعطى نفس الإجابة التى يعطيها ٥ × ٤ ؟

(أ) ٤ + ٤ + ٤ + ٤ + ٤

(ب) ٥ + ٥ + ٥

(جـ) ٥ + ٤

(د) ٥ ÷ ٤

٢ – ٣/٤ تساوى

(أ) $\frac{٦}{٨}$ (ب) $\frac{٥}{٨}$ (جـ) $\frac{٤}{٨}$ (د) $\frac{٣}{٨}$

ومن أهم الأهداف التعليمية التى تحتاج إلى تقويمها فى اختبارات الحساب هى القدرة على تطبيق معرفة المبادئ والأساليب أو الإجراءات (وتقاس باختبارات المسائل الحسابية من مختلف الأنواع).

٣– تقويم التحصيل فى العلوم:

مع زيادة التأكيد على العلم فى منهج المدرسة الابتدائية ككل، نجد من الطبيعى أن تسعى المدرسة الابتدائية إلى زيادة مدى ما يتضمنه برنامجها من العلوم، والأهداف التى نسعى إلى تحقيقها يجب أن تكون أكثر من مجرد معرفة حقائق معينة أو مصطلحات أو حفظ مبادئ وقوانين، وإنما يجب أن تتضمن قدرات مثل القدرة على إجراء تجارب بسيطة، واشتقاق المبادئ والتعميمات من البيانات أو إعطاء أمثلة للمبادئ المختلفة وتطبيق مثل هذه المبادىء فى شرح النتائج أو التنبؤ بها كما يجب أن تتعلق الأهداف باتجاهات وميول معينة.

وحيث أن أغلب معلمى المدرسة الإبتدائية يعوزهم الإعداد السليم فى العلوم فإنهم يحتاجون إلى المساعدة فى تحديد أنماط التعلم التى تنتج عن مقرر العلوم تحديداً واضحاً وفى الحصول على أمثلة خاصة من القدرات المختلفة، ومن الأساليب المفيدة فى هذا الصدد إعداد وانتقاء اختبارات جيدة وأسئلة جيدة فى الاختبارات، فكثير من القدرات يمكن التدليل عليها بطريقة جيدة عن طريق تطبيقها فى مواقف تعد أسئلة الاختبار أمثلة لها أو مواقف تتحول إلى أسئلة اختبار، ومعرفة الطريقة التى يمكن بها

تقويم هذه القدرات يساعد فى توضيح المعنى الحقيقى للقدرة عند المعلم وتشجعه على تدريسها.

وفيما يلى مثال:

الغرض العام	الهدف المرحلى	الهدف النوعى
أولا: على المواطن المتعلم أن يقدر مصادر الأمة.	أ. أن يفهم التلميذ أهمية المصادر الطبيعية مثل الماء والأرض والنبات والحيوان.	١ – أن يستطيع التلميذ أن يذكر أسباب ضرورة المحافظة على مصادرنا من الماء والأرض والنبات والحيوان. ٢ – أن يستطيع المتعلم تحديد الطرق المختلفة للمحافظة على المصادر الطبيعية.
ثانيا: يقيس المواطن المتعلم التقدم العلمى بإسهامه فى الرفاهية العامة.	ب. أن يشعر التلميذ ببعض الإسهامات الهامة التى يقدمها العلم فى مجال الرفاهية العامة.	٣ – أن يستطيع التلميذ أن يتعرف على إسهامات علماء مثل باستير واديسون وسولك ومركونى.
ثالثا: تتوفر لدى المواطن المتعلم وسائل دفاع ضد الدعاية.	جـ. أن يفهم التلميذ الطريقة العلمية فى حل المشكلات.	٤ – أن يستطيع التلميذ أن يحل مشكلات العلم بالوصول إلى النتائج من الأدلة والشواهد.
	د. أن يصل التلميذ إلى قرارات على أساس الأدلة والشواهد سواء فى مسائل العلم أو المسائل غير المتعلقة بالعلوم.	٥ – عندما يطلب من التلميذ أن يتخذ قرارات فى مواقف فرضية (فى مجالات العلوم وغيرها) فإنه يصل إلى تلك التى تدعمها الشواهد والأدلة.

أن الهدفين النوعيين المشتقين من الغرض العام الأول يهتمان بمعرفة المبادئ والطرق، ويمكن تقويم الهدف الأول بمجرد سؤال المتعلم أن يذكر قائمة ـ إما شفويا أو كتابة ـ بعدد معين من الأسباب التى تتعلق بالمحافظة على كل مصدر من المصادر الطبيعية المذكورة.

ويمكن للهدف الثانى أن يصاغ عل نحو يشبه الهدف الأول، وبالتالى يمكن تقويمه بأن تطلب من المتعلم أن يذكر قائمة شفوية، ويمكن أيضًا تقويم هذا الهدف باستخدام الاختبارات الموضوعية.

مثال: حتى يستطيع الفلاح أن يحصل من أرضه على محصول وفير عليه أن يزرع فى حقله.

(أ) نفس المحصول عاما بعد عام.

(ب) محصولا واحدا لثلاث سنوات متوالية ثم محصولا آخر للسنوات الثلاث التالية.

(جـ) محصولا واحدا لخمس سنوات ثم يليه محصول آخر للسنوات الخمس التالية.

ويمكن صياغة هدف رقم (١) بحيث يسمح للمتعلمين بانتقاء الأسباب المهمة للمحافظة على المصادر الطبيعية، وعلى هذا النحو يمكن تقويمه باستخدام أسئلة الاختيار من متعدد، أما الهدف النوعى المعتمد على الغرض العام الثانى فيمكن اختباره باستخدام أسئلة بسيطة مثل أسئلة المزاوجة بين أنواع الإسهامات التى قام بها مختلف العلماء.

والأهداف المشتقة من الغرض العام الثالث أصعب فى تقويمهما نوعا ما. ويمكن أن تفيد فى ذلك الاختبارات الشفوية واختبارات المقال، حيث يعرض على المتعلم مواقف مشكلة معينة ثم يطلب منه إعطاء الاستنتاجات والحلول المناسبة، كما يمكن تقويم هذه الأهداف باستخدام أسئلة الاختيار من متعدد من النوع التفسيرى كما هو مبين فى المثال الآتى:

مثال: قام أربعة متعلمين بالاشتراك فى تجربة لاستنبات الفول كما يلى:
قام المعتز بمسح بعض حبوب الفول بقطعة من القطن ثم تركها جافة.
قامت مها بمسح بعض حبوب الفول بقطعة من القطن ثم تركتها مبللة
قام خالد بوضع بعض حبوب الفول فى طين جاف وتركها جافة.
قام أحمد بوضع بعض حبوب الفول فى طين مبلل وتركه مبللا.

ومن نتائج هذه التجربة حاول المتعلمين فى الفصل الحكم على أى النتائج الاتية حول الصحيح:

(أ) تنبت حبوب الفول إذا تركت فوق قطعة من القطن الجاف.
(ب) تنبت الحبوب إذا تركت فوق قطعة من القطن المبلل.
(جـ) تنبت الحبوب إذا تركت فى الطين الجاف.
(د) تنبت الحبوب إذا تركت فى الطين المبلل.
(هـ) تحتاج الحبوب إلى أن تظل مبللة حتى تنبت.
(و) تحتاج الحبوب إلى أن تظل جافة حتى تنبت.
(ز) تحتاج الحبوب إلى التربة والماء حتى تنبت.

وبالنسبة لكل من العبارات الآتية افترض أنها تصف النتائج الحقيقية للتجربة وعليك أن تقرر أى الاستنتاجات السابقة هو الأقرب إلى الصواب إذا كانت النتيجة واحدة من هذه العبارات، حدد اختيارك بكتابة الحروف أ، ب، جـ، د، هـ، و، ز على الخط القصير أمام العبارات الخاصة بالنتيجة:

١ – حبوب المعتز ومها أنبتت ، ولم تنبت حبوب الآخرين...
٢ – حبوب أحمد أنبتت ، ولم تنبت حبوب الآخرين...
٣ – حبوب مها أنبتت ، ولم تنبت حبوب الآخرين...
٤ – حبوب مها وأحمد أنبتت ، ولم تنبت حبوب الآخرين...
٥ – حبوب خالد وأحمد أنبتت ، ولم تنبت حبوب الآخرين...
٦ – حبوب المعتز وخالد أنبتت ، ولم تنبت حبوب الآخرين...

وبهذه الطريقة يمكن قياس بعض جوانب القدرة على التفكير العلمى فى طفل المرحلة الابتدائية.

الفصل الثاني عشر

الاختبارات المرجعة إلى المحك

مقدمة :

يغلب على الاختبارات المنشورة أنها اختبارات معيارية، وأسباب غلبة هذا النوع من الاختبارات المنشورة المرجعة إلى المعيار أنها مصممة ومعدة وموضوعة بواسطة آخرين من غير مستعمليها، وإنما مبنية بحيث تكون قيمتها الكبرى فى كيف يتباين أداء جماعات مختلفة عليها.

وكان من الممكن ، مع هذا، أن تفسر درجات بعض الإختبارات فى ضوء عدد من العناصر التى يجيب عليها المتعلمين إجابة صحيحة بصرف النظر عن مقارنة هذا العدد بأداء الجماعة، وفى مثل هذه الحالة نقول أن الاختبار يمثل محكا أدائيا per-formance criterion ، ونقول أن المتعلمين الذين يجيبون على كل عنصر من عناصر الاختبار إجابة صحيحة، أو يجيبون إجابة صحيحة من عدد محدد سلفا من العناصر فى الاختبار. هؤلا المتعلمين يمكن أن نقول أنهم استوفوا حق الأداء الكامل المطلوب من الاختبار، هذا الاتجاه هو الذى يسمى الترجيع إلى المحك criterion-referenced أو الترجيع المحكى criterion-referencing.

ويجب ملاحظة أننا لانستطيع أن نسمى أى اختبار اختبارا مرجعا إلى المحك لمجرد أنه ليس اختبارا مرجعا إلى المعيار، ذلك لأن معنى الترجيع إلى المحك أن الأداء على الاختبار مربوط أو متصل بمراجع السلوك (*) behavior referents وأن يكون الاختبار مصمما ومنشأ على هذا الأساس (*) ، أى أن الاختبار بحكم طريقة تصميمه، يجب أن يمدنا بمعلومات عن قدرة المتعلم على القيام بأداءات معينة بصورة مطلقة in absolute terms فإذا كان الاختبار اختبار جمع وطرح الكسور مثلا، يجب أن يكون لدينا أساس ما لزعمنا أن المتعلم يعرف كيف يجمع ويطرح الكسور، إذا استطاع أن يجيب عن عناصر الاختبار كلها، أما إذا لم يستطع أى متعلم أن يتم الإجابة، وإذا كان للاختبار حقا محك أدائى سلوكى، فإنه من الممكن أن نخلص إلى أنه ما من متعلم استطاع أن يحقق محك الكفاءة proficiency criterion للمهارة المقاسة.

(*) النطاق: الأداءات التى يتطلبها المرمى، والتى منها نختار بنود الاختبار.

مصادر المحك:

ولكن من أين يأتى المحك الذى يرجع إليه الاختبار؟

يقترح كوكس وفارجاس (١٩٦٦) Cox & Vargos أن محكا رئيسيا للترجيع هو أن التدريب يؤدى إلى زيادة الكفاءة، فإذا أعطى اختبار لجماعة ما قبل التدريب وبعده، وكانت نتائج الجماعة بعد التدريب أفضل منها قبله فإن هذا يعنى أن الاختبار حساس لنواتج التدريب، أى أنه من الممكن تفسير نتائج الاختبار على أنها تدل على اكتساب المهارات والمعرفة التى قصد التدريب إلى إنتاجها، مثل هذا الاختبار يكون اختبارا مرجعا إلى المحك.

وأنها من المؤسف حقا أن يكون عدد قليل جدا من الاختبارات التى ينشئها المدرسون اختبارات مرجعة إلى المحك يمكن أن تتصف بالقوة والمتانة. ذلك أنه حتى يمكن اعتبار اختبار ما مرجعا إلى المحك، من الضرورى أن نحاول أن نحدد العلاقة بين الأداء على هذا الاختبار وسلوك محكى criterion behavior . وحتى يكون لدينا تبرير سليم لاعتبار اختبار ما اختبارا مرجعا إلى المحك، فإننا لابد أن نتبع الخطوات التالية فى بنائه وترجيعه:

١ – يعد تخطيط محتوى content outline يذكر فيه المهارات والمعرفة التى يعتبر الاختبارمحاولة لقياسها.

٢ – يميز الأداءات، أى المرامى القابلة للقياس measurable objectives التى سوف يكون المفحوص قادرا عليها، مع التسليم بأن هذا المفحوص قد اكتسب كفاءة فى المهارات والمعرفة المقاسة بالاختبار.

٣ – يميز المجال أو النطاق domain الذى يحدده كل مرمى ويكتب العناصر أوالبنود فى الاختبار وفق التفاصيل المعينة فى ذلك المجال أوالنطاق، ثم يختار عشوائيا عنصرين لكل مرمى ليبنى منها الاختبار.

٤ – يتأكد من صحة حقيقة أن المهارات والمعرفة المقاسة بالاختبار هى فى الواقع لازمة perquisite للمرامى الأدائية أو المرامى السلوكية المميزة المطلوبة فى الخطوة رقم ٢، وقد يكون هذا هو أكثر جوانب هذه العملية تغيرا، ذلك لأن عملية

إثبات صدق الاختبار validation تبدأ بحكم الشخص الذاتى، أى بإجراء صدق ظاهرى، ثم تمتد لتتضمن حكم مجموعة من الخبراء أو تتضمن بيانات أو معطيات حقيقية حصلنا عليها من إعطاء الاختبار لمجموعة أثبتت أداءها الكفء لترى ما إذا كانت لدى أفرادها المهارات والمعرفة الخاصة بالاختبار.

٥ – يقرر أو يحدد درجة قطع (أو درجة قاطعة) cutoff score أو درجة محك criterion score تعبر عن الأداء على الاختبار الذى يجب أن يحصل عليه المختبر ليظهر أنه قد أحرز كفاءة كافية فى المهارات والمعرفة ليكون قادراً على أداء ألوان السلوك المحكية criterion behaviors.

ملامح الاختبارات مرجعية المحك

إن أهم ملامح الاختبارات المرجعة إلى المحك هى:

١ – أنها مؤسسة على عدد من المرامى الأدائية أو السلوكية والتى تعتبر هذه الاختبارات محاولة لقياسها.

٢ – وأنها مصممة بحيث يكون لها درجة عالية من الملاءمة appropriateness لأنها قائمة على مرام، ومحتواها يرتبط بمراميها.

٣ – أنها تمثل عينات من السلوك أو الأداء الواقعى الحقيقى.

٤ – أن الأداء عليها يمكن أن يفسر فى ضوء درجات قاطعة محددة سلفا.

وأنه من الأهمية بمكان أن نوضح أنه ليس هناك فواصل قاطعة بين الاختبارات المرجعة إلى المعيار والاختبارات المرجعة إلى المحك. وفى الواقع من الممكن أن نفسر اختبارا ما متجهين أحد الاتجاهين أو كليهما على شرط أن تكون المرامى أو تصنيفات المحتوى المتميزة هى الأساس الذى استعمل فى كتابة عناصر الاختبار، على أن الفرق الأساسى بين الاختبارات المرجعة الى المحك والاختبارات المرجعة إلى المعيار هو أن الأولى مؤسسة على درجة قاطعة محددة سابقا، مفترضا أنها داخلة فى تصميم الأداء نفسه intrinsic أو هى وثيقة الاتصال بذلك الأداء، بينما النوع الأخير من الاختبارات مؤسس على أداء معيار، أى أن تفسير درجة الاختبار أساسه خارجى extrinsic.

إنشاؤها واستعمالاتها

إن الترجيع إلى المحك أكثر مناسبة فى تطبيقه فى مجال اختبار المهارة وأشد صعوبة فى تطبيقه لقياس السلوك المعقد مثل التفكير وحل المشكلة، كما يذهب (Ebel, 1970a, Cronbach, 1977) . وعلى الرغم من هذا فإن ثمة اهتماما وشغفا باستخدام الترجيع إلى المحك استخداما أوسع فى قياس التحصيل المدرسى، سواء بالنسبة إلى الاختبارات التى ينشئها المدرس، أو الاختبارات المنشورة.

ويستخدم الترجيع إلى المحك عن طريق:

١ – وضع، واختبار، مجموعة من المرامى تمثل النواتج المرغوب فيها للتدريس.

٢ – تصميم عناصر أو بنود، أو اختبارها بحيث نقيس كل مرمى قياسا دقيقا، أى تكون هذه العناصر أو البنود ممثلة لنطاق المرمى objective domain.

٣ – التحديد المسبق presetting لمستويات أداء مقبولة.

٤ – إعطاء الاختبار لمجموعة من المتعلمين وتقويم أدائهم فى ضوء عدد من المرامى التى يستطيعون أداء ما تتطلبه من أداء بنجاح.

ويتضمن الترجيع إلى المحك مهمة كتابة أو وضع مرام وعناصر اختبار لقياس هذه المرامى، وبينما نجد هذا يسيرا على المدرسين فى فصولهم، فإن كتابة مرام وعناصر لتستخدم على نطاق أوسع، لمنطقة أومحافظة مثلا، أمر على درجة كبيرة من التعقيد، وربما كانت مؤسسة اختبارية أكثر صلاحية وقدرة على القيام بها من المدرس، إلا أنه حتى المؤسسة الاختبارية ذاتها سوف تجد من الشاق أن تعد اختبارا مرجعاإلى المحك على أساس قومى أو حتى إقليمى، ذلك لأن مثل هذا الاتجاه سوف يضحى بتصويب الاختبار نحو حاجات نظام مدرسى بعينه، وإذا كانت الحقيقة هى أن المدخلات المحلية، أكثر ضرورة للاستخدام المحلى، وأن عددا جد قليل من الأفراد المحليين، المدرسين فى المدرسة مثلا، لديهم الوقت، أو المهارة لإنشاء اختبارات يمكن استخدامها فى نطاق واسع، هذه الحقيقة تحد كثيرا من الامكان والقيمة التطبيقية للاتجاه إلى بناء اختبارات تحصيلية مرجعة إلى المحك.

وعلى أية حال، فى الولايات المتحدة، نجد فى زيادة «بنوك» أو «مخازن» مرامى

الأداء performance objectives، وعناصر الاختبار وبنوده، وزيادة رغبة مديرى المدارس فى أن يكونوا قواعم مرامى الأداء لمناطقهم، ونمو شركات الاختبار فى هذا المجال الخاص بالاختبارات المرجعة إلى المحك، كل هذا زاد من احتمال أن الاختبارات المرجعة إلى المحك سوف تكون ميسرة للاستعمال فى قياس التحصيل المدرسى.

ولهذه الاختبارات المحكية، فيما يبدو، مزية السماح لكل منطقة تعليمية بأن توجه برنامجها التقويمى نحو أهدافها الخاصة، وأن تعرف مدى إحراز هدفها بشكل مطلق وليس بشكل نسبى، فليس كافيا أن تقول أن «محمدا» تعلم أكثر من «حامد»، ذلك لأنه مازال علينا أن نعرف عدد المرامى التى حققها كل منهما حتى تسمح بالتقدم نحو أهداف أخرى أكثر تركيبا، وسوف يقدم ناشرو الاختبارات مساعدة كبيرة للمدارس إذا أمدوها بعناصر الاختبار التى تتفق مع المرامى، حيث تكون كل مدرسة قادرة على أن تشكل اختبارها التحصيلى وتصوغه، موجها نحو حاجات متعلميها.

٥ – تفسير درجات الاختبار المرجع إلى المحك:

درجة الاختبار ليست سوى رقم ، ماذا تحمل هذه الدرجة من معنى ودلالة؟ ماذا تكشف عنه؟ هل أبدى المتعلم كفاءة فى المرامى المقاسة؟ هل لديه القدرة ليتحرك إلى مستوى أداء تال؟ هل لديه القدرة على أداء مهارات على مستوى مطلوب لبداية العمل فى مهنة ما؟

درجة الاختبار فى ذاتها لاتدلنا إلا على القليل جدا، وحتى تكون مفيدة يجب أن يعطى لها تفسير.

وإذا كان الاختبار قائما على عدد من المرامى، وكانت هذه المرامى ذاتها لها درجة من الصدق، فإن الدرجة على الاختبار سوف تخبرنا بشىء عن قدرة المتعلم على الأداء فى مجال الاختبار. ويجب أن يمدنا الاختبار المرجع إلى المحك بمعلومات عن درجة كفاءة المتعلم فى المرامى التى يقيسها الاختبار، فالدرجة المرتفعة على الاختبار يجب أن تدل على إحراز المرامى وتحصيلها، بينما الدرجة المنخفضة تدل على نقص التحصيل، وقد تظن أنه ربما كانت عناصر الاختبار سهلة جدا أو ضعيفة جداً، ومن ثم فإنها تكون غير منصفة فى قياس التحصيل. ولكن إذا كان اختبارك قد

استوفى محكات الملاءمة، والصدق، والثبات فإنه من الأرجح أن تكون عناصره مناسبة لقياس التحصيل، ومن ثم يمكنك أن تحكم بأن النجاح فى الاختبار يقيس إحراز المرامى وتحصيلها.

وحتى يكون تفسير درجة الاختبار التحصيلى دقيقا وسليما، ربما يكون من الأفضل أن تقيد نفسك بالدرجة الكلية على الاختبار، ولما كان الاختبار المرجع إلى المحك أكثر شبها بمجموعة من الاختبارات المصغرة Mini-tests يقيس كل منها مرمى واحدا، وكل منها يحتوى عنصرين، وأحيانا أكثر من عنصرين، لما كان هذا شأن الاختبار المرجع إلى المحك، فإن التحصيل يمكن أن يوصف عن طريق ذكر المرامى الخاصة التى أظهر فيها كل متعلم كفاءة بحكم نجاحه فى الاختبار المصغر المناسب أوالمتصل بهذه المرامى، وعلى هذا، فإنه لابد من أن تحتاج إلى أن تقرر ما هى الدرجة المرتفعة التى يمكننا اعتبارها مرتفعة. يمكن أن تسجل لكل متعلم اسم المرمى الذى أظهر فيه كفاءة، ولا بأس - إذا كان هذا مرغوبا - بإعطاء المتعلم ما يدل على مستوى الكفاءة الذى أحرزه فعلا. فتفسيرية مثل هذا الاختبار قائمة أساسا لا على الدرجة الكلية بل على دلائل درجة الكفاءة التى أمكن إحرازها لكل مرمى من المرامى التى يقيسها الاختبار، ولما كانت المرامى تمثل اهتمامنا الحقيقى، وكانت الاختبارات المرجعة إلى المحك تقيس تحصيل هذه المرامى فإننا نستطيع أن نركز اختبارنا وتفسيراتنا على ذلك التعلم الذى تريد أن تدفعه وترفعه.

الاختبارات التحصيلية المرجعة إلى المعيار

إن الترجيع إلى المعيار Norm- referencing واحد من أهم مكونات الإختبارات المقننة. والإختبارات التحصيلية المرجعة إلى المعيار تفسر درجاتها بمقارنتها بدرجات حصل عليها بالفعل تلاميذ آخرون. أولئك الذين يكونون الجماعة المعيارية أو جماعية المعيار norming group وتقرر ما إذا كانت درجة ما منخفضة أو مرتفعة بمقارنتها بدرجات حصل عليها متعلمون من مستوى الصف ذاته فى الجماعة المعيارية.

ويمكن القول أن كل اختبار ينشئه المدرس هو بالإمكان اختبار مرجع إلى المحك، لأنه مصمم لقياس درجة الكفاءة المنخفضة على عدد معين من المرامى، ومن العجيب أن كثيرا من المدرسين لايبنون اختباراتهم حول مرام معينة ومحددة، بل إن

كثيرين منهم يصححون اختباراتهم كما لو كانت مرجعة إلى المعيار بأن يقيمو تصحيحهم على أساس الأداء النسبى للمفحوص، فى حين أنه يجب تصحيح الاختبار فى ضوء أداء على مرام أو عناصر ، أى على أساس الترجيع إلى المحك.

الاهتمام الرئيسى فى الاختبارات المرجعة إلى المعيار هو: إلى أى مدى يمكن أن يقارن فرد أو مجموعة أفراد مع آخرين؟ والإجابات عن هذا السؤال مفيدة جداً فى ضمان تحقيق مستوى أدنى من الأداء النسبى فى الفصل، أو المدرسة، أو المنطقة التعليمية.

أما الاختبارات المرجعة إلى المحك فتسأل : بأى شكل يسلك فرد أو مجموعة من الأفراد؟ ما الذى يعرفه الفرد ذاته؟ ما الذى تعرفه المجموعة؟

ومن ثم فإن الترجيع إلى المعيار يعطى نتائج تعيينية، summative أى أنها تخبرنا أين يقف الفرد أومجموعة الأفراد، بينما الترجيع إلى المحك يعطينا نتائج تشكيلية formative ، أى أنها تخبرنا فى أى المجالات يمكن أن يوجه التدريس بحيث يتيسر تحصيل الكفاءة، الاختبارات المرجعة إلى المحك تساعد المدرسة على تعرف monitor تقدم المتعلم وتشخيص مواضع القوة والضعف ، ووصف التدريس وتوجيهه، وخاصة العلاجى منه.

وبينما يستطيع المدرسون إنشاء اختباراتهم التحصيلية المرجعة إلى المحك، فإن الاختبارات المنشورة تعطيهم درجة أرفع من ضبط نوع أداة الاختبار، وإمكانات أفضل فى ملامح تقرير نتائج الاختبار، وفى إمكان المدرسين أن يفسروا هذه النتائج فى ضوء مراميهم التدريسية الخاصة.

ومن الممكن بناء اختبار تحصيلى مقنن، معيارى، يكون فى الوقت ذاته اختبارا مرجعا إلى المحك، كيف؟

مقارنة بين الاختبارات المرجعة إلى المعيار، والاختبارات المرجعة إلى المحك

تعرف أن الاختبارات التحصيلية المقننة يتوافر لها ثلاث خصائص هى:

١ – أن عناصرها وبنودها جربت وحللت، وروجعت.

٢ – انتشارها واتساع استعمالها وإعادة استعمالها المقنن.

٣ – وجود معايير واستعمال هذه المعايير فى تفسيرها.

وأنه من الممكن بناء اختبار تحصيلى مقنن يكون فى الوقت ذاته مرجعا إلى المحك، وذلك بتغيير واحدة من الخصائص الثلاث السابقة، تلك هى الخاصية رقم (٣)، بحيث تكون هكذا.

«وجود مرام للتفسير واستعمالها، وتتضمن هذه المرامى محكات لتقويم الكفاءة، فعندئذ نكون قد وضعنا الأساس، على الأقل من حيث التعريف، لاختبار تحصيل «مقنن»، يكون اختبارا تحصيليا مرجعا إلى المحك.

هذا، وأن الفرق الوظيفى بين اختبار مرجع إلى المحك (رك CR) Criterion Referencedوآخر مقنن أومرجع إلى المعيار (ر ع NR) ذو ثلاث شعب:

الأولى: أن الاختبار المرجع إلى المحك موجه توجيها خاصًا إلى مرام بينما الاختبار المرجع إلى المعيار تكون درجاته أكثر كلية more global وشمولا، ومن ثم فإنك تحصل على درجات أكثر فى الاختبار المرجع إلى المحك، (لأنك تحصل على درجة واحدة لكل مرمى)، ولكن يوجد فيه عدد أقل من العناصر لكل درجة.

والثانية: درجات الاختبار المرجع إلى المعيار يمكن أن تفسر عن طريق المعايير التى يطعيها ناشروه، بينما درجات الاختبار المرجع إلى المحك تتطلب منك أن تضع أنت نقطة أو درجة القطع التى تدل على الكفاءة المناسبة.

والثالثة: عناصر الاختبارالمرجع إلى المعيارتكتب بحيث تنتج أقصى قدر من التنوع فى الأداء بين المتعلمين، بينما الاختبار المرجع إلى المحك تكتب عناصره بحيث تمثل نطاق المرامى كما يراه كاتب العنصر.

بعد أن نتمعن فى هذه الفروق بين الاختبار المعيارى NR (رع) والاختبار المحكى CR (رك) وبعد معرفة أن عناصر الاختبارالمعيارى قد كتبت مستعملة تخطيط محتوى content outline أو تصنيف مهارات skills classification ، فمن الممكن أن نفكر فى استخدام اختبار معيارى كما لو كان اختبار محك على شرط أن يذكر ناشره نتائجه بمجال المحتوى أومجال المهارة، وقد بدأ كبار ناشرى الاختبارات منذ الآن يفعلون هذا بالنسبة للاختبارات المعيارية أو المقننة مثل اختبار أيوا Iowa Test أو Metropolitan Achievement Tests of Basic Skills..

وأن تفسيرات نتائج الاختبارات التحصيلية المرجعة إلى المحك يغلب أن يكون مساعدا للمدرسين بصورة أكبر من تفسيرات الاختبارات المرجعة إلى المعيار، إلا أننا يجب أن ننتبه إلى نقطتين:

الأولى: من الضرورى أن تكون المهارات والمحتويات المقاسة على وجه الخصوص هى تلك التى تحتاج معلومات عنها، وأنها يوجد عدد كاف من العناصر لكل مهارة، وذلك لتوفير الثبات.

والأخرى: لما كانت العناصر مكتوبة لأقصى تمييز بين الدرجات المرتفعة والدرجات المنخفضة، فإنك قد تحتاج إلى فحص درجات متعلمينك على الاختبار لتحديد أى نقطة قطع، أودرجة قطع، تستخدمها للكفاءة المناسبة بالنسبة إلى كل مهارة، بدلا من استخدام نقطة قطع ثابتة مقررة، وبمعنى ما فإن هذا مازال صورة من صور الترجيع إلى المعيار، ذلك لأن اختبار نقط القطع سوف يقوم على الأداء النسبى relatìve performance للمتعلمين، مثلا، عدد العناصر المجابة صوابا من٧٠٪ من المتعلمين.

مراجع الباب الخامس

١ – أحمد زكى صالح: الأسس النفسية للتعليم الثانوى، دار النهضة العربية، ١٩٧٢ .

٢ – رايتسون، ج. واين وآخرون (ترجمة محمد محمد عاشور وآخرين): التقويم فى التربية الحديثة، مكتبة الأنجلو المصرية، ١٩٦٥ .

٣ – رمزية الغريب: التقويم والقياس النفسى والتربوى، الأنجلو المصرية، ١٩٧٠ .

٤ – سامى إبراهيم على: تقنين امتحانات البيانو، رسالة ماجستير المعهد العالى للتربية الموسيقية، ١٩٧١ .

٥ – فان دالين، ديوبولد (ترجمة محمد نبيل نوفل وآخرين): مناهج البحث فى التربية وعلم النفس: مكتبة الأنجلو المصرية، ١٩٦٩ .

٦ – فؤاد أبو حطب: دورالتربية فى تنمية التفكير الابتكارى، مجلة الفكر المعاصر، يونيه ١٩٧٠ .

٦ أ– فؤاد أبو حطب: الإمتحان الشفوى ماله وما عليه. ورقة ألقيت فى الندوة التى عقدتها وحدة التقويم الجامعى ١٧–١٨ ابريل ,٢٠٠٠ منشورة فى المجلة المصرية للدراسات النفسية ١١٩–١٢٤ – العدد ٣١ ابريل ٢٠٠١ .

٧ – ميجر، روبرت ف. (ترجمة جابر عبد الحميد جابر وسعد عبد العزيز نادر): الأهداف التربوية، بغداد، مطبعة العامى، ١٩٦٧ .

٨ – نعيم عطية: التقييم التربوى الهادف، بيروت، منشورات دار الكتاب اللبنانى، ١٩٧٠ .

9. Bloom. B. Taxonomy of Educational Objectives. London : Longmans, 1956.

10. De Cecco, J.P. The Psychology of Learning and Instruction. Englewood Cliffs, N.J., 1968.

11. Mager, R.F. Preparing Objectives for Programmed Instruction. Palo Alto, Calif. : Varian Associates, 1961.

12. Ebel, B.L. Maximizing Test validity in Fixed Time limits. Educ, psychol. Measurements, 1953, 8, 347-357.

13. Ebel, R.L. Measuring Educational Achievement, Prentice- Hall, 1965.

14. Gagne, R.M. Conditions of Learning, Holt., Rinchart, Winston, 1965.

15. Gronlund, N.E. Measurement and Evaluation in Teaching,Macmillan, 1965.

16. Linguist, E.F. (ed.) Educational Measurement. Amer. Council on Educ., Washington, D.C. 1951.

17. LindvaII, C.M. Testing and Evaluation : An Introduction, Harcourt, Brace, 1961.

18. Mager, R.F., McCann, J. Learner-Controlled Instruction Palo Alto, Clif Varian Associates, 1961.

19. Thorndike, R.L. (ed.) Educational Measurement (2nd ed.) Amer. Council ,on Educ., D.C., 1971.

الباب السادس
مقاييس الشخصية

الفصل الثالث عشر

وسائل التقرير الذاتى

مقدمة :

أشرنا فيما سبق إلى تقسيم المقاييس النفسية بصفة عامة إلى قسمين رئيسيين هما: مقاييس الأداء المميز، ومقاييس أقصى الأداء، وعرفنا أن فى النوع الأول من هذه المقاييس يكون الهدف من القياس هو معرفة ما يقوم به الشخص الآخر عادة، بينما النوع الآخر من المقاييس يكون الهدف هو معرفة أقصى ما يمكن أن يقوم به الشخص فى ظروف استثارة ودافعية غير عادية.

ويمكن أن نضع تحت القسم الأول، أى مقاييس الأداء المميز، اختبارات الشخصية، واختبارات الميول والاتجاهات بينما تدخل اختبارات القدرة بأنواعها المختلفة تحت القسم الثانى وهو اختبارات أقصى الأداء، وفى هذا الفصل نتكلم عن قياس الشخصية باعتباره قياسا للسلوك أو الأداء المميز، وسوف تتبع، بصفة عامة، التقسيم الذى حدده كرونباك (٢٧) لأنواع مقاييس الشخصية، فندرجها تحت هذه الأقسام الثلاثة:

(أ) قياس الشخصية عن طريق أحكام الآخرين والملاحظات المنظمة Judg-ments and systematic observations ، وهى ما سيتناوله الفصل الرابع عشر.

(ب) قياس الشخصية عن طريق اختبارات الأداء Performance tests وهى ما سيتناوله الفصل الخامس عشر.

(جـ) قياس الشخصية عن طريق التقرير الذاتى self-report وهى ما سيتناوله الفصل الحالى.

وقياس الشخصية عن طريق التقرير الذاتى Self-report على الرغم من أن تعبير السلوك المميز الذى تكررت الإشارة إليه يحمل طابعا «سلوكيا» فإن المذهب السلوكى لم يكن هو الذى أسهم فى نشأة وتطوير مقاييس الشخصية وإنما يرجع الفضل فى هذا إلى الاتجاه الظاهراتى (الفينومونولوجى phenomenology) ، فهذا الاتجاه يتميز بأنه يركز على العالم، كما يدركه الشخص، أو على العالم الذاتى للشخص، بما فى ذلك إدراك الفرد لذاته، وعلى أن السلوك إنما يتحدد ، ولايمكن أن يفهم إلا فى

ضوء هذا العالم الذاتى الداخلى، كما يذهب هذا الاتجاه إلى أن أغلب مشكلات التوافق التى يتعرض لها الفرد ترجع إلى إدراك الشخص للأحداث والأشخاص وليس لواقع هذه الأحداث أو الأشخاص فى ذاتهم، وبالتالى فإن العلاج النفسى، وفق هذا الاتجاه يجب أن يركز على تصحيح إدراك الشخص.

نشير إلى المدرسة الفينومونولوجية فى بداية كلامنا عن قياس الشخصية لأن أول مقاييس أو استفتاءات للشخصية كانت قد وضعت بهدف قياس هذا العالم الداخلى الذاتى من الإدراك والشعور، ففى العقد التاسع من القرن التاسع عشر وضع فرانسيس جالتون Francis Galton هذه المقاييس عندما إحتاج إلى طريقة يمكن أن يطبقها على عدد كبير من مفحوصيه أثناء دراساته للصور العقلية، ثم استعمل هذا النوع من الاستفتاءات بعد ذلك ستانلى هول Stanely Hall فى دراساته الموسعة عن نمو المراهقين ولكن الاستفتاء كانت له وظيفة خاصة بالنسبة إلى كل من جالتون وستانلى هول، فبالنسبة إلى جالتون كان التقرير الذاتى يستعمل على أنه الوسيلة الوحيدة الممكنة للحصول على معلومات عن أمور وأحداث فى عقل المفحوص، فى داخل رأسه، بينما كان التقرير الذاتى بالنسبة إلى ستانلى هول يستعمل بهدف تجنب الجهد الشاق والوقت الطويل الذى يمكن أن يستهلك فى الملاحظة المباشرة لسلوك المراهقين (٢٧ : ٤٦٤).

أما أول استفتاء وضع بقصد قياس الشخص فقد كان قائمة وود ورث للبيانات الشخصية Woodworth Personal Data Sheet وظهرت خلال الحرب العالمية الأولى عندما ظهرت الحاجة إلى تمييز أولئك الجنود الذين عندهم قابلية للإنهيار أثناء القتال، ولم يكن يتوافر عدد كاف من المتخصصين لتحقيق هذا الهدف، لهذا وضع وود ورث قائمة من الأعراض التى يهتم الفاحص بالتعرف عليها لاستبعاد من تظهر عندهم، وعندما أعطيت القائمة ميزت بالفعل أولئك الذين قرروا وجود عدد كبير من الأعراض عندهم ثم خضعوا لاختبار أكثر دقة، وكانت لاختبار البيانات الشخصية هذه قيمة كبيرة فى ذلك المجال.

ويعتبر اختبار وود روث الجد الأكبر لعدد كبير من استفتاءات التوافق، وهى التى تتكون أساسا من قوائم من المشكلات أو الأعراض أو الشكاوى التى على المفحوص أن

يختارها، إلا أن مثل هذه القوائم لاتدعى أنها تعطى وصفا دقيقا عن الشخصية وهى غالبا ما تعطى درجة واحدة تمثل مستوى التوافق. وهى تتكون من عناصر يمكن أن تميز بين أشخاص معروف أنهم غير متوافقين وبين آخرين معروف أنهم عاديون، والاستعمال الرئيسى لمثل هذه المقاييس هو التعرف على أولئك الأشخاص الذين يحتاجون إلى رعاية وإرشاد نفسى، وعلى هذا يمكن أن ننظر إليها على أنها أدوات فرز screening تساعد فى تحديد الأشخاص الذين يختارون أعراضا وصفات وانتقادات تتصل بهم شخصيا، وقد عدله بيرت وعربه واستعمله أحمد زكى صالح (١ : ٨٥١ - ٨٥٢).

ولايمكن اعتبار استفتاءات التوافق مقاييس دقيقة لسمات محددة، فالمعلومات التى تمدنا بها معلومات ليس لها صفة العمق، لهذا فإنه فى الفترة بين ١٩٢٠ - ١٩٤٥ تغيرت النظرة إلى مثل هذه الاستفتاءات وذلك بتأثير المدرسة السلوكية واستنكارها للتأمل الباطنى الذاتى الذى هو أساس اختبارات التوافق، فيجب أن يكون المقياس بديلا لملاحظة السلوك ، وعلى هذا يجب أن تؤكد أسئلته بشكل أكبر على ما يفعله الشخص لا على ما يشعر به أو يفكر فيه، وعلى هذا اتسعت مجالات الاختبارات بحيث تصف أكبر عدد ممكن من جوانب الشخصية، وتلخص الاستجابات بإعطاء درجات لعدد من السمات traits أو لأنماط الاستجابات، وذلك لأنه كان ينظر إلى الشخصية فى ذلك الوقت على أنها مجموعة أو حزمة من العادات، وكان يحكم على الشخص حسب قوة سماته مثل: الثقة بالنفس، الإصرار، سهولة تكوين صداقات... إلخ. والسمة القوية هى السمة التى تصف استجابة يقوم بها الشخص عادة وبكثرة.

وقد كانت قائمة السمات أو أقسام السلوك التى تعطى لها درجات فى الاختبارات تختار بشكل قطعى ودون خطة معينة فى الاختبارات الأولى، فمثلا مثل الثقة بالنفس كان مصدرها الخبرة اليومية، بينما تشتق سمة أخرى مثل الانطواء من نظرية من نظريات الشخصية ، وقد ظهرت عشرات من الاختبارات بهذا الشكل، تأخذ اللاحقة منها من السابقة عليها بصنعة عبارات أو أسئلة وتضيف إليها وتصممها بحيث توضع فى تجمعات سمات مختلفة، إلا أن أكثرها شيوعا هو اختبارا بيرنرويتر للشخصية Bernreuter Personality inventory.

وقد ترجمه إلى العربية محمد عثمان نجاتى(*) وتعطى فيه الدرجات للدلالة على: النزعة العصابية، الاكتفاء الذاتى، الانطواء، والسيطرة، ويتكون الاختبار من ٢٥ سؤالا يجيب عنها المفحوص بنعم أو لا. وقد وضعه بيرنرويتر على أساس أن سلوك الفرد فى موقف واحد قد يكشف عن سمات شخصية مختلفة، وعلى هذا يمكن أن ينظر إلى كل سؤال من أسئلة الاختبار على أنه يقيس أكثر من سمة من السمات، وبالتالى يمكن أن يستخدم الاختبار الواحد للكشف عن أكثر من سمة تبعا لطريقة معالجة الاستجابات التى نحصل عليها.

وقد كانت دراسة فلاناجان (٢٩) Flangan لاختبار بيرنرويتر بداية لاستعمال درجات السمات كما تحددها القواعد الإحصائية، لقد اعتقد فلانجان أن السمات لايمكن أن نطلق عليها أسماء خاصة بها إلا إذا كانت الارتباطات فيما بينها منخفضة، وعلى هذا قام بحساب معاملات الارتباط لدرجات اختبار بيرنرويتر من تطبيقه على ٣٠٥ مراهق، ووجد أن هذه السمات ليست مستقلة، فقد وجد أن الانطواء لايختلف إلا قليلا عن النزعة العصبية، وعندما طبق التحليل العاملى وجد عاملين هما: الثقة بالنفس والميل الاجتماعى يمكن أن يفسر البيانات التى نحصل عليها من قياس السمات الأربعة الأصلية، وأنشأ مفاتيح خاصة لقياس العاملين أو السمتين اللتين وصل إليهما أى الثقة بالنفس والميل الاجتماعى، وقد ارتبطت درجات هاتين السمتين ارتباطا ضعيفا جداً، وعلى هذا أصبح من الممكن أن نقول أنهما يمثلان جانبين مستقلين من التقرير الذاتى.

بعد هذا حدث تطور آخر له مغزاه، ذلك هو إخضاع نظريات الشخصية تماما لبحث إحصائى عن أبعاد dimension يمكن أن تلخص الشخصية. فقد أدت الارتباطات الداخلية بين عناصر الاختبارات إلى أن اقترح جيلفورد مثلا، أن الانطواء يمكن أن نقسمه إلى انطواء اجتماعى (ج)، وانطواء فكرى (ف)، واكتئاب (ك)، والتقلب المزاجى (ت)، وضبط النفس (ض)، وعلى هذا أنشأ الاختبار الذى أسماه: اختبار العوامل ج ـ ف ـ ك ـ ت ـ ض Inventory of Factors S-T-D-C-R.

(*) لاتتوافر لدينا بيانات سيكومترية عن هذا الاستخبار فى البيئة المصرية.

ثم أضاف بعد هذا ثمانية جوانب أخرى للشخصية، ولم تكن مقاييس جيلفورد مستقلة، ولهذا أعاد بعض الباحثين تنظيم هذه المقاييس فى أنماط تجعلها أكثر كفاءة، من هؤلاء مثلا ثرستون الذى استطاع أن يفسر كثيرا من المعلومات فى مقاييس جيلفورد الثلاثة عشرة بسبعة عوامل فقط وأعاد تسميتها بحيث أصبحت هكذا: التأمل، الميل الاجتماعى، الاتزان الانفعالى، الحيوية أو الذكورة، السيطرة، النشاط ، الاندفاع، وقد اقتبس منه أحمد زكى صالح اختبار الصفات الانفعالية.

وهكذا نجد أن واحدا من علماء النفس يصنف عناصر الاختبار التى تقيس السمات بشكل دقيق تفصيلى، بينما يأتى التالى ليضعها فى مجموعات صغيرة، ثم يأتى آخر فيعيد تقسيم المجموعات بطريقة جديدة، ويعطى كلا منها أسماء جديدة للعوامل التى يصل إليها، وبهذا لاتنتهى هذه اللعبة، كما يعبر كرونباك (٢٧ : ٤٦٧) عن رأيه أنه ما لم ترتبط قوائم السمات بنظرية محددة وبمحكات خارجية، فإن عملية الاختبار لايمكن إلا أن تكون عملية تفضيلية شخصية، وليس هناك اتفاق فى الوقت الحاضر بين أصحاب التحليل العاملى، بالنسبة إلى عدد العوامل التى تم تحديدها بشكل دقيق، أو أفضل تنظيم لها، أو أنسب الأسماء لها.

أما الاتجاه الأحدث فى قياس الشخصية بصفة عامة فهو تحديد المفاهيم والتكوينات النظرية Constructs على أساس نظرية للشخصية تم إعداد العناصر فى الاختبارات بصورة تؤدى إلى الحصول على معلومات عن هذه المفاهيم والتكوينات، وقد كان لنظرية يونج jung عن الشخصية تأثير فى الدراسات الأولى عن الانطواء، ولكن هذه النظرية لم تتعد اقتراح بعض عناصر اختبارات وضعت لمجرد المحاولة، أما الآن فهناك دراسات واسعة عن اختبار مايرز- بريجز Myers-Briggs Inventory والتى اشتقت عناصره ومفاتيح تصحيحه صراحة من نظرية يونج (٥١)، وتوجد مقاييس أخرى تمثل هذا الاتجاه منها مقياس التفضيل الشخصى لادوارد Edwards Personal Preference Schedule(EPPS) . والذى يشتق من نظرية موري Mur-ray عن الحاجات الخمس عشرة، ومن أمثلة هذه الحاجات: الإنجاز، والخضوع، والنظام، والعرض، والاستقلال الذاتى، السيطرة، التأمل النفسى (المقصود بهذه الحاجة تحليل دوافع ومشاعر الفرد الذاتية وعند الاخرين)، ويتكون المقياس من ٢١٠ زوجا من العبارات، وعلى المفحوص أن يختار عبارة من كل زوج على أنها أكثر تمثيلا

لخصائص شخصيته، ومن أمثلة هذه الأزواج.

(أ) أحب أن أتحدث عن نفسى مع الآخرين.
(ب) أحب أن أعمل لهدف معين حددته بنفسى.

(أ) أشعر بالاكتئاب عندما أفشل فى أمر ما.
(ب) أشعر بالاضطراب عندما أتحدث إلى جماعة.

ويهدف هذا المقياس إلى تقدير القوة النسبية للخمسة عشر متغيراً المرتبطة بالحاجات.ويمكن تقويم الخمسة عشر درجة المرتبطة بالحاجات على أساس معايير مئوية وتائية لطلاب وطالبات الجامعة، وهذه المعايير تقوم على عدد ٧٤٩ طالبا، و٧٦٠ طالبة موزعين على جميع أنحاء الولايات المتحدة، وقد أضيفت معايير مئوية للكبار ينتمون إلى بيئات حضرية وريفية.

أما عن ثبات المقياس فقد ذكر دليل الاختبار معاملات ثبات لمقاييسه الفرعية الخمس عشرة معتمدة على إعادة الاختبار بعد أسبوع واحد تتراوح بين ٠,٧٤ ، ٠,٨٨ ، بينما معاملات الثبات المعتمدة على التجزئة النصفية تتراوح بين ٠,٦٠ ، ٠,٨٧ .

أما بالنسبة إلى صدق مقياس ادواردز، فعلى الرغم من قلة البيانات الواردة عنها فى دليل الاختبار ذاته، فإن أبحاثا مستقلة حول صدقه قد نشرت، إلا أنه أعطت نتائج متعارضة، وتذهب أنستازى (١٨ : ٤٥٤) إلى أن مقياس التفضيل الشخصى له مزايا كثيرة إلا أنه محتاج إلى أمرين:

أولا: مراجعة للتخلص من بعض جوانب الضعف الفنية فيه، وبصفة خاصة فيما يتصل بشكل العناصر وتفسير الدرجة، والثانى، دراسات صدق تجرى بدقة مستفيدة من أساليب تحليل أنماط الدرجات مناسبة لدرجات هذا المقياس التى لا تتحدد فيها قوة الحاجة بصورة مطلقة ولكن بالنسبة إلى قوة الحاجات الأخرى عند الفرد، وعلى هذا فإن الإطار المرجعى فى هذا النوع من الدرجات هو الفرد ذاته وليس عينة التقنين: وقد يكون شخصان متفقان فى درجاتهم على المقياس إلا أنهما يختلفان اختلافا جوهريا فى القوة المطلقة لحاجتهم.

وقد عرب هذا المقياس جابر عبد الحميد (١٧ : ٤٧٨ – ٤٨٠) وحسب ثبات الاختباربطريقة التجزئة النصفية، وذلك بتطبيقه على طلاب الجامعة، وكان يتراوح بين ٠,٣٤ ، ٠,٧٧ وقد استعمل هذا المقياس فى بعض الدراسات المقارنة بين الثقافات وغيرها.

ومن المقاييس الأخرى التى اعتمدت على نظرية مقياس تيلور للقلق الظاهر Tayler Manifest Anxiety Scale وقد ارتبط فى تصميمه بالبحث فى نظرية هل- سبنس Hull-Spence السلوكية، وسوف نعود إليه بعد دراستنا لاختبار مينسوتا المتعدد الأوجه الذى نوليه اهتماما خاصاً.

اختبار مينسوتا المتعدد الأوجه للشخصية:

Minnesota Multiphasic Personality Inventory (MMPI)

يعتبر هذا الاختبار من أهم الأدوات التى وضعت لقياس الشخصية عن طريق التقرير الذاتى، لهذا سوف نتناوله بشىء من التفصيل، نجد لهذا الاختبار مكانا بين اختبارات الشخصية تشبه مكان اختبار سترونج Storng بين اختبارات الميول، ذلك لأنه يشارك سترونج فى الأساس التجريبى الذى صمم وفقا له، وكذلك يشبهه من حيث تعرضه لبحث عميق وطويل من جانب واضعيه، وقد ظهر هذا الاختبار فى وقت كانت الحاجة إليه شديدة، وكان الاعتماد عليه كبيرا أثناء وبعد الحرب العالمية الثانية التى شهدت توسعا كبيرا فى علم النفس المرضى والعلاج النفسى، أى أن نشأته كانت أصلا نشأة إكلينيكية وهذا مما ساعد على انتشاره، وكذلك ساعد على انتشاره أيضاً بعض الأمور الفنية المبتكرة فيه ومنذ نشره فى عام ١٩٤٠ ونشر دليله فى عام ١٩٤٣ فإنه استثار فيضا من البحوث التى استخدمته.

وقد كان تصميم الاختبار فى الأصل على يد عالم نفسى هو ستارك هاثاوى Starke Hathaway وطبيب نفسى هو ج. س. ماكنلى J. C. Mckinley (٣٥) وكان الهدف من هذا التصميم هو المساعدة فى تشخيص المرضى النفسيين وتحليل السمات التى يميز الشذوذ النفسى الذى يؤدى إلى العجز.

يتكون الاختبار من ٥٥٠ عبارة يسأل المفحوص أن يصنفها وفق أحد أقسام ثلاثة، أى يجيب عنها إما بنعم أو لا، ولا أستطيع أن أجزم أو لا أدرى وفى الصورة الفردية للاختبار

كانت العبارات تطبع على بطاقات منفصلة وعلى المفحوص أن يصنفها فى ثلاث مجموعات، وبعد ذلك طبعت العبارات فى كتيب وتسجل الاستجابات فى ورقة الإجابة. وقد صممت الصورتان الفردية والجماعية للكبار أى من ١٦ سنة وما بعدها، إلا أنهما استخدما بنجاح مع مراهقين أصغر من هذا السن، وبصفة عامة يفضل استخدام الصورة الفردية وخاصة عند الأشخاص المضطربين أو من ذوى المستوى التعليمى أو العقلى المنخفض.

وتتنوع عناصر الاختبار المتعدد الأوجه للشخصية وتتسع من حيث محتواها فتشمل مجالات مثل: الصحة، والأعراض الجسمية ـ النفسية، والاضطرابات العصبية، والاتجاهات نحو الجنس، والسياسة، والدين، والمسائل التربوية والمهنية والعائلية، كما تشمل كثيرا من المظاهر والأعراض العصابية والذهانية المعروفة مثل المخاوف المرضية، والاضطهاد، ونزعات تعذيب الذات وتعذيب الاخرين، وغيرها.

وعندما نشر الاختبار فى صورته الأولى كان يعطى درجات على تسعة «مقاييس كلينيكية»، ويتكون كل مقياس منها من عناصر ميزت بين جماعة مرضية محدودة وبين جماعة ضابطة عادية، وعلى هذا فإن هذه المقاييس قد أنشئت بطريقة تجريبية، وذلك عن طريق تقسيم العناصر وفق محك معين وهو التشخيص النفسى المعروف ووفق هذه الطريقة أعدت المقاييس التالية:

١ – توهم المرض ـ هـ س	يسمى ثلاثى العصاب
٢ – الاكتئاب ـ د	
٣ – الهستريا ـ هـ ى	
٤ – الانحراف السيكوباتى ـ ب د	
٥ – الذكورة ـ الأنوثة ـ م ف	
٦ – البارانويا ـ ب أ	
٧ – السكيائينيات ـ ب ت.	
٨ – الفصام ـ س ك	
٩ – الهوس الخفيف ـ م أ	
صفر– الإنطواء الاجتماعي*	

1. HS: Hypochondriasis
2. D : Depression
3. Hy: Hysteria.
4. Pd: Psychopathic deviation
5. Mf: Masculanity-femininity
6. Pa: Paranoia
7. Pt: Psychasthenia
8. Sc: Schizophrenia
9. Ma: Hypomania
0. Si: Soial Introversion

(اناستازى ١٩٩٧ : وأوروبينا: ٣٥٣)

(*) أضيف المقياس العاشر فى الطبعة الثانية mmpi-2 من المقياس الأصلى mmpi-1 وقد اشتقت من أداء مجموعتين متضادتين فى المرحلة الجامعية تم اختيارهم بناء على درجاتهم المتطرفة في سمتي الإنطواء – الإنبساط.

وفيما يلى تعريف مختصر بما يدل عليه كل من هذه المقاييس (١٥ : ١٥٢ – ١٦٥).

١ – توهم المرض: وهومقياس للاهتمام الزائد بالوظائف الجسمية، والقلق الذى لا مبرر على الصحة.

٢ – الإكتئاب: ويرتبط هذا المقياس بالشعور بالانقباض والتشاؤم والإحساس بعدم الأهمية الذاتية.

٣ – الهستيريا: ويقيس هذا المقياس درجة تشابه المفحوص بالمرضى الذين تظهر عندهم أعراض تتخذ صورة شكاوى عامة أو محددة مثل العمى أو الشلل أو القىء أو الاضطرابات المعوية، أو الأعراض القلبية.

٤ – الانحراف السيكوباتى: وهو مقياس لدرجة تشابه المفحوص بجماعة السيكوباتيين الذين تتمثل الصعوبة الرئيسية عندهم فى نقص الاستجابة الانفعالية العميقة وقلة الاهتمام بمشاعر الآخرين، وعدم المبالاة بالمعايير الاجتماعية.

٥ – ـ الأنوثة: وهو مقياس للذكورة والأنوثة فى أنماط الاهتمامات. وقد استخرجت عبارات المقياس نتيجة مقارنة استجابات الذكورة من ذوى الاهتمامات الذكرية باستجابات الذكور من ذوى الاهتمامات الأنثوية، أما الفروق بين استجابات الذكور والأناث قد درسها واضعا الاختبار إلا أن الفروق لعبت دوراً صغيرا نسبيا فى هذا المقياس وتدل الدرجات المرتفعة على هذا المقياس على غلبة أنماط الاهتمامات الخاصة بالجنس الآخر.

٦ – البارانويا: هذا مقياس للتشكك الزائد، والحساسية المفرطة وهواجس الاضطهاد.

٧ – السيكاثينيا: وهو مقياس للمخاوف المرضية والسلوك القهرى، الذى قد يكون صريح مثل التكرار الزائد لغسل اليدين، أو ضمنيا كما يتمثل فى عدم القدرة على التحرر أو الهروب من فكرة متسلطة.

٨ – الفصام: ويكشف هذا المقياس عن التفكير أوالسلوك الخاطىء الشاذ والذى لا يتفق مع الاتصال بالواقع وفهمه.

٩ – الهوس الخفيف: وهو مقياس للنشاط الزائد فى الفكر والعمل، وعدم التركيز الطويل، والانشغال بعدد كبير من الأمور والمشروعات فى وقت واحد.

وتعطى بيانات جماعة التقنين من العاديين درجة معيارية (ب) يمكن تحويل للدرجات الخام إليها بحيث تخطط النتائج على بطاقات صحيفة نفسية profile sheet وتتركز الأهمية الأولى على الدرجات التى تزيد عن ٧٠ والدرجة ٥٠ هى (متوسط جماعة المرجع) . وهذه الدرجة الفاصلة، أى ٧٠، هى درجة مختارة على أساس قطعى اعتسافى، أى ليس لها أو تبرير نظرى، وعلى أى حال نجد أن مفسرى نتائج الاختبار يميلون إلى الاهتمام بكل الارتفاعات، أو القمم، فى بطاقة الصحيفة النفسية، سواء كانت هذه الارتفاعات تتخطى خط الـ ٧٠ درجة أو لا تتخطاه.

مقاييس الضبط: ومن العلامات المميزة والخاصة بمقياس مينسوتا المتعدد الأوجه استعماله لأربعة مقاييس تسمى الصدق، والصدق هنا معناه الضبط Control لأنها فى الواقع تمثل ضوابط للإهمال، وسوء الفهم، والتحليل، ولتأثير أنماط الاستجابات والاتجاهات نحو أخذ الاختبارات، وهذه المقاييس هى:

١ – الأول؟، وتمثل درجة هذا المقياس مجموع العناصر التى أجاب عنها المفحوص ب «لا أستطيع أن أجزم»، ومن المرغوب فيه أن تكون هذه الدرجة أقل ما يمكن.

٢ – الثانى ل L، ويمثل الكذب، ويعتمد على مجموعة من العناصر التى تجعل المفحوص يبدو فى شكل مقبول اجتماعيا، ولكن ليس من المحتمل أن تكون الإجابة عنها صادقة فى الاتجاه المرغوب فيه، ومن أمثلة هذا العنصر.

لا أقول الصدق دائما.

وعلى الرغم من أن الإجابة الصحيحة المعتادة تكون نعم، إلا أن الإجابة المقبولة اجتماعيا تكون لا، ومن المفترض أن الشخص الذى يريد أن يظهر نفسه فى صورة مقبولة يحصل على درجة مرتفعة فى المقياس ل عن طريق تحريف استجاباته لعبارات المقياس، وتدل الدرجات العالية على هذا المقياس على أن الإجابات لايمكن الثقة بها، ولكنها لا تدل على أن المفحوص يتعمد الكذب.

٣ – الثالث ف F، وهو مقياس الخطأ، ويتكون من العناصر التى تدل أنها أجيبت من مجموعة التقنين فى نفس الاتجاه الذى أجاب فيها المفحوص، على الرغم من أن هذه العناصر لايمكن أن تمثل نمطا معينا من الشذوذ النفسى فإنها تعبر عن سلوك غير مرغوب فيه، وعلى هذا فإنه من غير المحتمل فى الواقع أن يعرب مفحوص واحد عن كل أو معظم هذه الأعراض وترتفع الدرجة على هذا المقياس عندما لا يستطيع المفحوص أن يعطى إجابة مميزة لسبب من الأسباب كأن لا يستطيع القراءة أو الفهم بدرجة مناسبة، أو الإهمال، سواء عن قصد أو غير قصد، كما قد يرجع ارتفاع الدرجة إلى الخطأ فى التصحيح، وبصفة عامة كلما زادت الدرجة على هذا المقياس زاد احتمال أن بعض العوامل تدخلت لتقلل من صدق الصحيفة النفسية (أو البروفيل).

٤ – الرابع ك K ويمثل مقياس التصحيح، وفى هذا المقياس استعمال لتركيبه أخرى من العناصر المختارة، وتعطى درجة هذا المقياس دلالة على الاتجاه نحو أخذ الاختبار، وهى متصلة بكلا المقياسين السابقين ل. ف. ولكن يعتقد أنها أكثر خفاء منهما وتدل الدرجة العالية على هذا المقياس على اتجاه دفاعى defensiveness أو على محاولة المخادعة فى الاتجاه الحسن، «Fake good» أو التحريف المقصود نحو الطرف السوى، كما تدل الدرجة المنخفضة على الصراحة المفرطة، ونقد الذات، أو المحاولة المتعمدة للمخادعة فى الاتجاه السيء fake bad أو الميل إلى إظهار النفس بالمظهر غير السوى.

ومن المهم أن نشير إلى أن استعمال مقاييس الضبط الأربعة هذه ليس مقننا تماما ولكنه يترك جزئيا لحكم مستعمل الاختبار، وكذلك نشير إلى أن هذه المقاييس تخضع باستمرار لإعادة التحديد والمراجعة فى ضوء البحوث الجديدة.

كانت تلك هى المقاييس الإكلينيكية التسعة الأساسية ثم المقاييس الأربعة الأخرى الخاصة بضبط الاختبار، إلا أنه منذ نشر الاختبار أنشئ عدد كبير جداً من المقاييس الجديدة بواسطة باحثين آخرين لم يشاركوا فى وضع المقاييس الأصلية، من هذه المقاييس الجديدة مقياس عن الانطواء الاجتماعى (س ى)، Social Introversion (Si) وقد أضيف إلى المقاييس التسعة الأصلية وهو مضمن فى الاختبار ورقمه (صفر)، ويقيس هذا المقياس النزعة إلى الانزواء والابتعاد عن الاتصال الاجتماعى بالآخرين.

وبهذا فإن الاختبار فى صورته الحالية يعطى ١٤ درجة على الأربعة عشر مقياس منها التسعة الأصلية، ومقياس الانطواء الاجتماعى، ثم مقاييس الضبط الأربعة.

ومعايير المجموعة الضابطة الأصلية والتى تتكون من حوالى ٧٠٠ فردا مذكورة فى صورة درجات تائية (ت)، أودرجات معيارية بمتوسط ٥٠ وانحراف معيارى قدره ١٠، وتستعمل هذه الدرجات المعيارية فى رسم الصحيفة النفسية، وأى درجة تعادل ٧٠ أو تزيد عنها، أى تقع فوق المتوسط بدرجتين انحراف معيارى، ينظر إليها بصفة عامة على أنها تميز انحرافا مرضيا ومما يجب ملاحظته أن المعنى الإكلينيكى لنفس الدرجة قد يختلف من مقياس إلى مقياس آخر، فمثلا الدرجة ٧٥ على مقياس توهم المرض وعلى مقياس الفصام لا تدل على نفس الدرجة من حدة الانحراف.

أما تفسير الدرجات المنخفضة على الصحيفة النفسية للاختبار فليس على درجة كافية من الوضوح ، إذ يعتقد الاكلينيكيون أن وقوع درجات أقل من٥٠ كثيرا قد يكون له مغزى تشخيصى، ولكن لم توضع بعد تفسيرات منظمة لهذا.

على أى حال، فإن هناك قدرا من نتائج البحوث التى أجريت على هذا الاختبار يوضح أنه بصفة عامة كلما زاد عدد وحجم الدرجات المنحرفة على الاختبار، زاد احتمال أن يكون الشخص مصابا باضطراب شديد.

ومن الواضح أن الاستعمال الرئيسى للاختبار المتعدد الأوجه هو التشخيص المتمايز أو الفارق differential diagnosis عندما يستعمل فى هذا المجال لابد أن يمارس مستعمله منتهى الحذر ولايجأ إلى التفسير الحرفى للمقاييس الاكلينيكية فيه، لأن الأمر أكثر تعقيدا من تلك الأسماء المعطاة للمقاييس الأصلية، فليس معنى الدرجة المرتفعة عل مقياس الفصام مثلا أن الشخص مصاب بالفصام، فقد يحصل شخص عادى على نفس الدرجة، بينما قد يحصل المصابون بالفصام فعلا على درجات مرتفعة على مقاييس أخرى، ولكى يستبعد احتمال سوء التفسير للدرجات عل المقياس الواحد فإن أرقاما من صفر إلى ٩ حلت محل أسماء المقاييس فى طبعات أخيرة من الاختبار، ولكى يسهل النظر فى نمط الدرجات أو الصحيفة وفحصها وضع نظام ترميز عددى للصحيفة النفسية. وفى هذا النظام يدل تتابع وترتيب أرقام المقاييس،

بنظرة سريعة، على النقط المرتفعة والمنخفضة فى الصحيفة النفسية لشخص ما، فمثلا الرمز ٩٤ – ٢ يدل على درجة مرتفعة على المقياس رقم ٤ (ب د) وعلى المقياس رقم ٩ (م ١) بترتيب متناقص، أى أن الدرجة على المقياس الأول أعلى منها على المقياس الثانى، كما يدل هذا الرمز على درجة منخفضة على المقياس رقم ٢(د).

والواقع أن أسماء المقاييس التى استبعدت كانت تستند إلى تقسيم إكلينيكى يقوم على أساس نظرى مشكوك فيه على الرغم من شيوعه وقت وضع الاختبار، فالتقسيمات المستعملة والتى وضعت لها مقاييس الاختبار كانت محل شك ومراجعة فى علم النفس المرضى أما فى المجال العلمى الواقعى فإن هذه التقسيمات لم تعد مقبولة.

ومن دواعى الحذر فى تفسير درجات المقاييس أن بعض هذه المقاييس لا تتمتع بثبات مرتفع، وتعتبر هذه النقطة ضعف فى الاختبار، لأن فعالية أى تحليل للصحيفة النفسية تقل بسبب أخطاء الصدفة فى الدرجات التى تعتمد عليها فى هذا التحليل، فإذا كانت درجات المقاييس المفردة غير ثابتة ومتداخلة أى تترابط فيما بينها ترابطا عاليا فإن كثيرا من الاختلافات فيما بين الدرجات ربما تكون ناتجة عن الصدفة.

وتتراوح معاملات ثبات الاختبار الناتجة عن إعادة الاختبار، بالنسبة إلى الراشدين الأسوياء وغير الأسوياء بين ٠,٥٠ و ٠,٩ وكانت الفترة بين الاختبار وإعادته تتراوح بين بضعة أيام إلى عام، إلا أن هناك دراسة أجريت على طلاب الجامعة كانت معاملات الثبات فيها عموما أقل من السابقة، كما أظهرت الدراسة نفسها تنوعا كبيرا بين مقاييس الاختبار كانت تتراوح بين ٠,٠٥ لمقياس ب ١، ٠,٨١ لمقياس ب ت.

وثمة عيب آخر فى الاختبار يتصل بمدى تمثيل عينة التقنين، فالدرجات المعيارية التى اشتقت منها تنظيمات الصحائف النفسية موضوعة فى ضوء جماعة ضابطة تتكون من ٧٠٠ راشد هم الذين اختبروا فى التقنين الأصلى للاختبار، وتبدو هذه العينة قاصرة جدا إذا قورنت بعينات التقنين الأخرى الأكثر شمولا والتى استعملت، مثلا فى تقنين اختبارات القدرة وقد أكدت أبحاث كثيرة ضرورة أن تكون

المعايير متنوعة تنوعا كبيراً فى جماعات أصل مختلفة، ومن هذه الأبحاث ما وضح أن متوسطات درجات طلبة الكليات كانت فوق ٥٠ بشكل مطرد على بعض المقاييس، كما بينت دراسة أخرى أن ٣٩٪ من طلبة الكليات حصلوا على درجات فوق ٧٠ على مقياس أو أكثر من مقاييس الاختبار.

ويتأثر الاختبار المتعدد الأوجه، مثل سائر اختبارات الشخصية، بالاختلافات الثقافية العامة بين المجتمعات بل وبالمستويات الثقافية الفرعية داخل المجتمع، وقد أظهرت دراسات كثيرة أجريت فى بلاد مختلفة ارتفاعا ملحوظا فى بعض مقاييس الاختبار عندما وضعت الصحائف النفسية على المعايير الأمريكية الأصلية، وأن تفسير مثل هذه الاختلافات فى ثقافة المجتمع أو ثقافاته الفرعية يتطلب معرفة بخصائص هذه الثقافات وظروفها، فقد يكون الاختلاف فى درجات الاختبار المتعدد الأوجه راجعا إلى اختلاف الأفراد فى تفسير عبارات الاختبار أو تعليمات الاختبار، فمثلا قد يرجع الارتفاع فى بعض المقاييس عند بعض الجماعات إلى اعتقادات متوارثة نحو التواضع وتقليل شأن الذات، وكذلك تتأثر الدرجة بالاختلافات الثقافية يتصل بالسلوك الذى يعتبر مقبولا اجتماعيا، إلا أن الدرجات المرتفعة على مقاييس هذا الاختبار، بالنسبة إلى بعض المجتمعات والجماعات، قد تدل على وجود مشكلات انفعالية جقيقية يمكن إرجاعها إلى أساليب تنشئة الطفل، أو الصراعات فى الأدوار الاجتماعية، أو غير ذلك من أساليب الاضطراب الانفعالى.

أما عن الترجمة العربية لاختبار الشخصية المتعدد الأوجه فقد أعدها عطية هنا، ومحمد عماد الدين إسماعيل، ولويس كامل مليكة، ووضعت له معايير من البيئة المحلية، وحسبت معاملات الثبات للمقاييس بطريقة التجزئة النصفية للمقياس ، وبطريقة كيودر- ريتشاردسون ، وقد استخدم الاختبار فى صورته العربية فى عدد من الدراسات (١٥: ٣١٦ – ٣٢٤).

اختبارات مشتقة من اختبار الشخصية المتعدد الأوجه:

إلى جانب حركة البحوث الضخمة التى حركها الاختبار المتعدد الأوجه فقد كان له فى ذاته خصوبة عالية من حيث اشتقاق مقاييس جديدة منه أو تأسيس اختبارات وتطويرها معتمدة عليه.

فمن المقاييس الجديدة الكثيرة التى استخرجت منه عن طريق استعمال مفاتيح تصحيح جديدة نذكر ما يلى:

١ – مقياس التعصب (Pr_) ويهدف إلى قياس العوامل السيكولوجية المرتبطة بالتعصب ضد جماعات أقلية.

٢ – مقياس المكانة الاجتماعية ـ الاقتصادية (St) ويقيس بعض العوامل السيكولوجية المرتبطة بالمكانة الاجتماعة ـ الاقتصادية للشخص.

٣ – مقياس الانحراف الجنسى.

٤ – مقياس قوة الأنا.

ويعتبر مقياس تايلور للقلق الظاهر (٥٩) الذى سبقت الإشارة إليه، من المقاييس التى اعتمدت على اختبار الشخصية المتعدد الأوجه، وقد استعمل هذا المقياس استعمالا واسعا فى البحث وفى الممارسة الاكلينيكية، وقد أنشىء أصلا للاستعمال فى التجارب لاختبار فروض معينة تتصل بتأثير الدافع على التعلم، وتحقيقا لهذا اختيرت عبارات معينة من الاختبار المتعدد الأوجه على أنها تدل على تعبير ظاهر عن القلق، وكانت نتيجة هذا الاختبار ٥٠ عبارة، زيد عليها عبارات إضافية لا تتصل بالقلق ومختارة من مقاييس الضبط الثلاثة ل، ف، ك. وقدترجمه إلى العربية الدكتوران مصطفى فهمى ومحمد أحمد غالى.

اختبار مينيسوتا للإرشاد النفسي

كذلك نتج عن الجهود التى بذلت لتعديل الاختبار المتعدد الأوجه بحيث يستعمل مع تلاميذ المرحلة الثانوية والطلبة الجدد فى الكليات أن أنشئ اختبار منسوتا للإرشاد النفسى Minnesota Counseling Inventory ويتكون هذا الاختبار من ٤١٢ عبارة على المفحوص أن يجيب عنها بصواب أو خطأ، ومعظم العبارات مأخوذة من اختبار الشخصية المتعددة الأوجه، ويعطى اختبار منسوتا للإرشاد النفسى درجات فى سبعة مجالات هى العلاقات الأسرية، والعلاقات الاجتماعية، والثبات الانفعالى، والمسايرة، والتكيف مع الواقع، والمزاج ، والقيادة، وعلى الرغم من أن مقاييس اختبار الإرشاد النفسى تسمى باسم الطرف المقبول، أى الثبات الانفعالى مثلا بدلا من الطرف الآخر

وهو الاضطراب الانفعالى، فإنها تتشابه مع مقاييس الاختبار المتعدد الأوجه المسماة بأسمائها المرضية، فمثلا مقياس «المسايرة» فى اختبار الإرشاد النفسى يشبه إلى حد كبير مقياس الانحراف السيكوباتى (Pd) فى الاختبار المتعدد الأوجه، وكذلك مقياس التكيف مع الواقعين يشبه كثيرا مقياس الفصام (Sc) ولهذا الإختبار مقياسان للضبط يشبهان مقاييس الضبط فى الاختبار المتعدد الأوجه وتشير بعض الدراسات (فى ٢٧ : ٤٩٦) إلى أن مقاييس هذا الاختبار ذات قدرة إيجابية ولكنها ضعيفة على تمييز التلاميذ الذين يعرف أنهم غير متوافقين أسريا من غيرهم ممن يتمتعون بتوافق جيد.

وللاختبار ثبات بإعادته بعد ثلاثة شهور يتراوح بين ٠,٧٠ ، ٠,٨٠ إلا أن تفسير هذا الاختبار يحتاج إلى حذر كبير، وقد نقل هذا الاختبار إلى اللغة العربية محمد عماد الدين إسماعيل وسيد عبد الحميد مرسى باسم مقياس الإرشاد النفسى، ولا تتوافر عنه بيانات سيكومترية فى البيئة المصرية.

ومن الاختبارات التى اعتمدت على الاختبار المتعدد الأوجه اختبار كاليفورنيا النفسى California Psychological Inventory فهو يشتق نصف عباراته من الاختبار المتعدد الأوجه ومن أمثلة مقاييسه: السيطرة ، النزعة الاجتماعية ضبط النفس، الإنجاز عن طريق المسايرة، والانجاز عن طريق الاستقلال والمرونة، وتعتبر أنستازى (١٨ : ٤٤٩) هذا الاختبار من أفضل اختبارات الشخصية المستعملة حاليا، وقد خضع لدراسات وتحسينات عديدة أعطته هذه المكانة، ولهذا الاختبار صورة لتلاميذ الثانوى هى اختبار كاليفورنيا للشخصية California Personality In-ventory نقلها إلى اللغة العربية جابر عبدالحميد ويوسف الشيخ، ولاتتوافر عنها بيانات سيكومترية فى اللغة العربية، كما توجد صورة من هذا الاختبار لتلاميذ

(*) أجرى عطية محمود هنا دراسة على «اختبارالشخصية للأطفال»، استعرض فيها صدقه وثباته فى البيئة المصرية. راجع:
عطية محمود هنا اختبار الشخصية للأطفال وقيمته فى البحوث النفسية، المجلة الاجتماعية القومية مايو ١٩٦٥..

المرحلة الابتدائية نقلها إلى اللغة العربية عطية محمود هنا باسم «اختبار الشخصية للأطفال»(*).

نماذج أخرى ممثلة من اختبارات الشخصية المعتمدة على التقرير الذاتى

عرضنا بشيء من التفصيل اختبار الشخصية المتعددة الأوجه لأنه «النموذج الأمثل» لاختبارات الشخصية التى تعتمد على التقرير الذاتى للمفحوص، إلا أن الاختبارات التى تنتمى إلى هذا النوع كثيرة ومتنوعة إلى درجة كاد يستحيل حصرها جميعا وخاصة فى هذا المجال المحدود، لهذا نختار نماذج ممثلة من هذه الاختبارات التى لم تناقش فى الصفحات القليلة الماضية مراعين أن يكون لكل منها ميزته التى تجعله جديرا بهذا العرض الموجز.

قائمة مونى للمشكلات Mooney Problem Check list

وقد وضعها روس مونى Ross L. Mooney وزملاؤه، ولها عدة صور من مراحل التعليم المتوسط حتى الجامعة والراشد، ويحدد فيها المفحوص مشكلاته فى أحد عشر مجالا منها الأخلاق والدين، النواحى المالية والمعيشية، والتوافق للعمل المدرسى، والعلاقات الاجتماعية، وتشير الدرجة المرتفعة إلى الحاجة إلى الإرشاد النفسى، وتتخذ العناصر التى يحددها المفحوص أساسا للمناقشة الفردية والجماعية وقد استخدمت هذه القائمة فى عدة دراسات عربية.

اختبار التوافق لـ بل The Bell Adjustment Inventory

منذ صدور هذا الاختبار فى عام ١٩٣٤ (٢٤) وهو يستعمل على نطاق واسع، والصورة الخاصة بالطالب مصممة للانتقاء السريع لطلاب المدارس الثانوية والكليات بهدف الإرشاد النفسى، وهذه الصورة تعطى درجات توافق فى مجالات أربعة هى: الأسرة، الصحة، الناحية الاجتماعية، والناحية العاطفية. والصورة الأخرى خاصة بالكبار، وهى أقل استعمالا، وقد زيد عليها مجال خامس هو مجال التوافق المهنى، وقد ترجم الصورة الخاصة بالطلاب محمد عثمان نجاتى، ولاتتوافر عنه بيانات سيكومترية فى البيئة المصرية.

مقياس ألبورت لدراسة السيطرة والخضوع Allport A-S Reaction Study

يعتبر هذا المقياس من أقوى اختبارات الشخصية المبكرة، ويوصف بأنه يقيس السيطرة ascendance والخضوع submission لأنه يقيس نزعة الشخص إلى أن يسود ويسيطر على زملائه، أو أن يساد ويسيطر عليه فى علاقات المواجهة اليومية،ويبدأ كل عنصر بوصف موجز لموقف من مواقف الحياة اليومية العادية فى اجتماع أو فى مدرسة أو سيارة عامة، ويذكر أسلوبين أو أربعة لمواجهة الموقف، وعلى المفحوص أن يحدد أى الأساليب يعبر أكثر عن سلوكه، وتختلف الاستجابات فى درجة السيطرة والخضوع التى تمثلها وتعطى لها أوزان مناسبة لهذه الدرجة فى التصحيح، وفضلا عن شيوع هذا الاختبار فإنه قدأثر فى كثير من الاختبارات التى أنشئت بعده ومن المهم أن نشير إلى أن السيطرة كانت إحدى السمات التى برزت بوضوح فى دراسات التحليل العاملى.

البروفيل الشخصى لجوردن Gordon Personal Profile

يعتبر هذا الاختبار، وقائمة الشخصية للمؤلف نفسه، من اختبارات الشخصية التى تعتمد على الاختيار الإجبارى forced - choice للتغلب على ما يسميه المرغوبية الاجتماعية، أى ميل المفحوص إلى أن يبدو مقبولا اجتماعيا، وكذلك للإقلال من الغموض وبعض الصعوبات الأخرى التى تصاحب عملية تصحيح اختبارات الشخصية.

ويقيس البروفيل أربعة جوانب للشخصية هى: السيطرة، والمسئولية، والاتزان الانفعالى، والاجتماعية (أو النزعة الاجتماعية) وهذه الجوانب مستقلة نسبيا، والبروفيل مناسب للاستخدام مع طلاب المدارس الثانوية والجامعة والراشدين.

ويتكون البروفيل من ١٨ مجموعة من العبارات الوصفية تشمل كل مجموعة على أربع عبارات وتمثل كل عبارة إحدى سمات الشخصية الأربع، المشار إليها من قبل، وتتشابه جملتان من الأربع من حيث أن لها قيمة تفضيلية عالية، أما الجملتان الأخريان فتتساويان فى القيمة التفضيلية المنخفضة، ويطلب من المفحوص أن يضع علامة على جملة من الجمل الأربعة على أساس أنها تشبهه إلى حد بعيد وعلى جملة أخرى على أساس أنها تشبهه إلى أقل درجة وبهذا لا يستطيع المفحوص أن يستجيب استجابة قبول لجميع العبارات كما يحدث فى كثير من مقاييس التقرير الذاتى الأخرى

أى أن البروفيل يعتبر أقل قابلية للتشويه من جانب أولئك الذين يحاولون أن يعطوا صورة طيبة عن أنفسهم.

وقد اقتبس البروفيل وأعده بالعربية جابر عبدالحميد، وفؤاد أبو حطب، ووضعا له كراسة تعليمات تتضمن وصفا له وتعليمات التصحيح كما وضعا له معايير مئوية، وأجريت دراسات عن معاملات المقاييس الأربعة للبروفيل وصدقه، كما استعمل فؤاد أبو حطب البروفيل فى دراسة (١٤) جاءت مدعمة لصدق التكوين الفرضى.

قائمة لشخصية لجوردون Gordon Personal Inventory

وتعطى هذه القائمة قياسا سريعا ومناسبا لأربع من سمات الشخصية هى: الحرص، والتفكير الأصيل، العلاقات الشخصية، الحيوية. ومن الممكن استخدامها فى المدارس الثانوية والجامعة ومع الراشدين، وتستخدم هذه القائمة، مثل البروفيل الشخصى، طريقة الاختيار الإجبارى، وهى تتكون من ٢٠ مجموعة من العبارات الوصفية، وتشتمل كل مجموعة منها على أربع عبارات، وتمثل كل عبارتين يعتبرهما الأشخاص العاديون متساويتين فى قيمة التفضيل العالية، وعبارتين يعتبرونهما متساويتين فى قيمة التفضيل المنخفضة، ويطلب من المفحوص أن يضع علامة أمام عبارة واحدة من العبارات الوصفية على أساس أنها تنطبق عليه أكثر من الصفات الثلاث الأخرى، كما يضع علامة أخرى أمام عبارة أخرى باعتبارها أقل انطباقا عليه من العبارتين الأخرتين، وتتميز القائمة بسهولة التطبيق والتصحيح والتفسير، وقد ترجمها إلى العربية فؤاد إبو حطب وجابر عبدالحميد جابر ووضعا لها كراسة تعليمات تتضمن وصفا لها وتعليمات التطبيق والتصحيح كما وضعت لها معايير مئوية، وحسبت لها معاملات الثبات والصدق، واستعملها فؤاد أبو حطب فى الدراسة التى سبقت الإشارة إليها (١٤).

وقد قامت آمال أحمد مختار صادق بتقنين كل من اختبار البروفيل الشخصى (١-أ) وقائمة الشخصية لجوردون (١-ب) على البيئة السعودية حيث طبق اختبار البروفيل الشخصى على عينات مختلفة من الطالبات السعوديات بالمدارس الثانوية (ن= ١٠٥) ومعاهد دور المعلمات (ن =٦٠)، ومن طالبات الجامعة (ن=٩٠)، ومن المعلمات بالمرحلة الابتدائية (ن =١٥). وقد استخدم فى حساب الصدق عدة طرق

(الصدق المرتبط بالمحكات حيث كان المحك مقياس تقدير للسمات الاربعة قامت بإعداده معدة الاختبار، حيث تم تطبيقه على ٤٢ طالبة يسهل التعرف عليهن وطلب من كل منهن أن تعطي تقديرا لذاتها فى السمات الأربع وتقدير «أفضل» زميلتين يتم اختيارهما بواسطة كل طالبة. وقد حسب معامل الارتباط بين هذا التقدير الذاتى والتقدير الموضوعى والمحدد بدرجة الاختبار، وقد جاءت المعاملات دالة فيما عدا سمة السيطرة، إلا أن المعاملات لم تكن دالة بمحك الاقران فى كل من سمة المسئولية والإتزان الإنفعالى، كذلك استخدم محك الاستبصار الذاتي ويتحدد اجرائيا بأنه الفرق الحسابى بين تقدير الذات وتقدير الاقران، وكانت أكبر هذه المعاملات فى سمة الإتزان الإنفعالى، كذلك استخدم محك تقدير الرؤساء بالنسبة لعينة المعلمات، إلا أن المعاملات لم تكن دالة. وفسرت هذه النتيجة إلى طبيعة التقدير الحر فى مقياس التقدير فى مقابل التقدير المقيد فى المقياس.

وقد استخدم أيضا صدق التكوين الفرضى والذى تم من خلاله اختبار فرضين الأول فى تحديد الفروق فى الأداء الاختبارى بين طالبات المرحلة الثانوية وطالبات الجامعة لصالح طالبات الجامعة،وقد تحقق الغرض بالنسبة لسمة المسئولية فقط.

أما بالنسبة لثبات الاختبار فقد تم تحديده بطريقة اعادة التطبيق بفاصل زمنى قدره شهران وتراوحت المعاملات بين ٠,٤٤ ، ٠,٦٣ وجميعها دالة.

وقد حسبت المعايير المئينية لكل من طالبات المرحلة الثانوية وطالبات دور المعلمات مستقلة عن معايير الطالبات بالجامعة نظرا لوجود فروق دالة بين أداء المجموعتين.

أما بالنسبة لاختبار قائمة الشخصية فقد تم تطبيقه على ذات العينات السابق الإشارة لها واتبعت نفس استراتيجية تحديد المعالم السيكومترية للإختبار حيث تم استخدام محك تقدير الذات وتقدير الاقران وتقدير الرؤساء وكانت جميع معاملات الارتباط دالة بين السمات المقاسة ومحك تقدير الذات فيما عدا سمة العلاقات الشخصية، أما بالنسبة لمحك تقدير الاقران كان المعامل الدال فقط هو الحيوية، ومحك الاستبصار الذاتى كانت جميع معاملات الارتباط دالة فيما عدة ايضا العلاقات الشخصية، أما محك تقدير الرؤساء فكانت المعاملات دالة فى سمتى الحرص والعلاقات الشخصية.

وقد حسب الثبات بإعادة التطبيق بفاصل زمنى شهرين وكانت جميع المعاملات دالة عنة مستوى ٠,٠١ كما حسب المعايير المئينية للعينات المختلة.

مقياس المسئولية الاجتماعية:

تأليف سيد أحمد عثمان، وهو مقياس لبعد من الأبعاد الاجتماعية ـ الأخلاقية للشخصية على أساس تصور وتحليل نظرى للمسئولية الاجتماعية متمثلة فى عناصرها الثلاثة وفى الاهتمام، والفهم، والمشاركة، وفى المستويات المتدرجة فى كل عنصر من هذه العناصر، كما حددها وناقشها مصمم المقياس (٥)، ويتكون المقياس من صورتين الصورة ت لتلاميذ الثانوى والصورة ك لطلاب الجامعة والراشدين.

وتتكون الصورة ت من المقياس من ١١٥ عنصرا، وهى عبارات تعكس ألوانا من السلوك أو الآراء، منها عبارات موجبة وأخرى سالبة ويطلب من المفحوص أن يحدد إجابته وفق مقياس متدرجة من أربع نقط.

وتسير الصورة ك من المقياس على النمط ذاته، وهى مقتبسة ومختارة من الصورة الأولى إلا أنها تتكون من ٨٥ عنصرا، منها أيضًا عناصر موجبة وأخرى سالبة.

وتتضمن كراسة تعليمات المقياس ، بصورتيه، شرحا للتحليل النظرى للمسئولية الاجتماعية، ثم بيانات عن وصف كل صورة من المقياس، وعن صدق الصورة (ت) وثباتها، وإشارة إلى الدراسات التى أجريت على هذه الصورة حتى الآن، ثم طريقة التصحيح بالنسبة إلى الصورتين، وفى كراسة التعليمات أيضًا بيان بإمكانات الاستعمالات المختلفة للمقياس فى صورتيه فى النواحى التربوية، وفى دراسة الشخصية، وفى العلاج النفسى، وفى الدراسات الاجتماعية.

والدراسات مستمرة فى إعداد معاييرالمقاييس واستكمال دراسات الثبات والصدق وخاصة للصورة ك وقد أجريت طوال العقد المنقضى منذ ظهور المقياس عدة دراسات عليه أكدت فعاليته، ويعتبر هذا المقياس أداة بارعة مرجوة تحتاج إلى عمل طويل ممتد متنوع تزداد معه الأداة دقة وكفاءة، كما يزداد معه التصور النظرى للمسئولية الاجتماعية اتساعا ووضوحا وعمقًا.

اختبارات كاتل فى الشخصية:

تعتمد هذه الاختبارات على نتائج البحوث العاملية التى أجراها ريموند كاتل R. Cattell ومعاونوه، وقد يكون أشهرها استخبار العوامل الستة عشر فى الشخصية

Sixteen Personality Factor Questionnaire (16 PF). الذى يقيس ١٦ سمة، كما أعدت قائمة أخرى للمراهقين (من ١٢ – ١٦) وتسمى Jr. - Sr High School Personality Questionnaire. وأعدها باللغة العربية سيد محمد غنيم وعبد السلام عبد الغفار باسم «استفتاء الشخصية للمرحلة الإعدادية والثانوية»، وتقيس ١٤ سمة. وكذلك أعدت قائمة ثالثة للمستويات العمرية من ٨ سنوات إلى ١٢ سنة وتسمى IPAT Children's Personality Questionnaire وأعدها باللغة العربية عبد السلام عبد الغفار وسيد محمد غنيم باسم استفتاء الشخصية للمرحلة الأولى». وقد حسبت المعايير التائية للاختبارين الأخيرين على عينات مصرية كما حسبت ثبات استفتاء الشخصية للمرحلة الأولى بطريقة الصورة المتكافئة (ن = ١٩٢ طفلا)، ولاستفتاء الشخصية للمرحلتين الإعدادية والثانوية بطريقة إعادة الاختبار (ن= ١٣٠ تلميذا). إلا أنه لاتتوافر بيانات سيكومترية عن صدقها فى البيئة المصرية.

قائمة إيزنك للشخصية:

هذا الاستخبار من إعداد عالم النفس الإنجليزى إيزنك وتقيس الانبساط مقابل الانطواء، والعصابية، ونقلها إلى اللغة العربية جابر عبد الحميد ومحمد فخر الإسلام، ولاتتوافر لدينا معلومات سيكومترية عنها فى البيئة المصرية.

استبيان مستوى الطموح:

هذا الاستبيان من تأليف كاميليا عبد الفتاح، ويتألف من ١٩ سؤالا مصنفة إلى ٧ فئات هى: النظرة للحياة، الاتجاه نحو التفوق، تحديد الأهداف والخطة، والميل إلى الكفاح، وتحمل المسئولية والاعتماد على النفس، والمثابرة، والرضا بالوضع الحاضر والإيمان بالحظ، وحسبت له المعايير المئينية (ن = ١٢٢ طالبة، ١٩٤ طالبا). وقد ذكرت المؤلفة معاملات ثبات الاستبيان وصدقه وتمييزه بين الأسوياء والعصابيين(*).

اختبار الإتجاهات التربوية:

يعد هذا الإختبار من الإختبارات المهمة فى مجال التربية والذى اعدته آمال

(*) كاميليا عبد الفتاح: كراسة تعليمات استبيان مستوى الطموح، النهضة المصرية، ١٩٧١.

صادق سنة ١٩٨٨ فى اطار نتائج البحوث التى قام بها ايزنك فى اطار نظريته المدعمة بنتائج البحوث العاملية طوال فترة تجاوزت النصف قرن والتى وصل فيها إلى عامل التحرر/ المحافظة والذى تأكد فى دراسات متعددة، إلا أن عامل الرقة/ الصرامة لم يحظ بنفس القدر من الإهتمام لأنه لا يتوافر له نفس النوع من المحكات الخارجية.

وقد اعتمد بناء المقياس على عامل الرقة/ الصرامة حيث أكدت نتائج البحوث ان العاملين على درجة عالية من الاستقلال، كما تبين ان بعدى الرقة / الصرامة ازاء المسائل التربوية أكثر الاتجاهات الاجتماعية قوة.

ويتألف الاختبار من ٦٠ عبارة يتم الإجابة عليها على تدرج خماسى وتدل الدرجة العالية على الرقة ازاء المسائل التربوية وتقيسها ٢٣ مفردة، ٣٧ مفردة تقيس قطب الصرامة وقد صيغت عبارات المقياس فى هيئة جمل خبرية ويتم معرفة مدى موافقة الطالب عليها أو معارضته لها.

وقد تم تقنين الإختبار على ٣٦٩ طالبة، ٢٣٧ طالبا من طلاب الجامعة، وعينة أخري من طلاب الدبلوم الخاصة والماجستير فى مصر، وكذلك تقنينه على طلاب وطالبات جامعيات من المملكة العربية السعودية ويوضح الجدول التالى الخصائص السيكومترية للإختبار، وجميع المعاملات التى تم الوصول لها دالة عند مستوى ا.و

جدول (٣٠)

المعالم السيكومترية لإختبار الاتجاهات التربوية*

عينات التقنين	الصدق	الثبات
طلاب كلية التربية الموسيقية	المرتبط بالمحكات ٠,٢٧٦	اعادة التطبيق ٠,٩١٨
دراسة شحاته محروس طه	المرتبط بالمحكات ٠,٤٢١	
	المرتبط بالمحكات ٠,٨٨٢	
طلاب دور المعلمين والمعلمات	المرتبط بالمحكات ٠,٦٤	التجزئة النصفية ٠,٨٦٨
دراسة فوقية عبد الفتاح	المرتبط بالمحكات ٠,٤٢	
طلاب وطالبات السعودية	المرتبط بالمحكات ٠,٨٤	التجزئة النصفية ٠,٨٩٨
دراسة آمال صادق	المرتبط بالمحكات ٠,٦٤	

أما معايير الاختبار فقد اعدت له معايير مئينية.

تقنين اختبار المخاوف للأطفال

أعد هذا الإختبار محمد عبد الظاهر الطيب ١٩٩٠، وقد قامت مها فؤاد أبو حطب بإعادة تقنينه فى دراستها المسحية عند الأطفال ٢٠٠٤ . وقد هدفت دراستها إلى تحديد أهم المخاوف المرضية الشائعة لدى أطفال المدارس الإبتدائية من سن ٨,٥- ١٢,٥ من غير المحولين من الجنسين ومن مستويات اقتصادية اجتماعية مختلفة (ن = ٢٣٦)

وقد استخدم الإختبار كأداة مسحية أولية لتحديد أهم حالات الرهاب لدي عينة الدراسة، ثم تلا ذلك استخدام المعايير التشخيصية للدليل التشخيصى الإحصائى الرابع DSM IV للإضطرابات النفسية للأنواع المختلفة للرهاب عند الأطفال.

وقد قامت الباحثة بتحديد صدق وثبات اختبار المخاوف المرضية للأطفال، فقد تم تحديد الثبات بطريقة التجزئة النصفية حيث بلغت قيمة معامل الإرتباط بين النصفين ٠,٧٥٩ وبعد التصحيح بمعادلة سبيرمان - براون من أثر التجزئة ٠,٨٦٣ وهو معامل دال عند مستوى ٠,٠١، كما تحددت قيمة معاملات الثبات لمتغيرات للإختبار من خلال تحديد اشتراكيات العوامل التى توصلت لها عند استخدام التحليل العاملى لتحديد صدق الإختبار حيث توصلت إلى ثلاثة عوامل. كما قامت الباحثة بتحديد مئينيات للإختبار.

وقد أثبتت الدراسة أن ٥,٧٠ من الأطفال يعانون من مخاوف مرضية فعلية، وأن حوالى ٢١,٧٦ ٪ يعتبروا فى المرحلة الحدية (Marginal) أى أكثر عرضة للإصابة بالرهاب - أي تتحول مخاوفهم إلى مخاوف مرضية. كما خلصت الدراسة إلى انتشار العديد من المخاوف المرضية خاصة الرهاب المحدد بأنواعه:

١- الرهاب المحدد المرتبط بموقف مثل : الخوف من الظلام - وسائل المواصلات العامة - الأماكن المغلقة - الأماكن المزدحمة.

٢- الرهاب المحدد المرتبط بالحيوانات مثل : الحيوانات الأليفة - الحشرات.

(*) لمزيد من التفصيل راجع دراسة التقنين فى المصدر (ا جـ).

٣- الرهاب المحدد بمواقف الحياة مثل : الخوف من الأماكن المرتفعة.

٤- الرهاب المحدد بأنواع أخرى مثل : الإصابة بالأمراض.

هذا بالاضافة إلى انتشار الرهاب الاجتماعى، وأن هناك أطفال يعانون من أكثر من نوع وان هناك اختلاف بين الذكور والإناث وأيضا يختلف أنواع الرهاب بين المستويات الإقتصادية الإجتماعية، كما أشارت الباحثة إلى ان انتشار الرهاب بين أطفال المدرسة الإبتدائية يعتبر ظاهرة تحتاج إلى مزيد من الوعى والإهتمام.

مقياس مفهوم الذات للأطفال *

إعداد فاروق عبد الفتاح موسى، فاتن فاروق عبد الفتاح موسى

Children's Self Concept Scale Ellin, V Piers & Daie B Harris**

أعد المقياس للبيئة المصرية فاروق عبد الفتاح موسى، فاتن فاروق عبد الفتاح موسى عن الأصل الأجنبى Chilfren's Self Concept Scaleوالذى ألفه إلن ف. بيرز وآخرون، وهو أداة سريعة لتقدير مفهوم الشخص عن ذاته (سن ١٠-١٧ سنة) ويستغرق تطبيقه ٢٠ دقيقة وهو سهل التطبيق إلا أن تفسير الدرجات يجب أن يتم بمتخصص فى علم النفس (قياس نفسى).

وقد إعتمد بناء المفردات فى أصله على العبارات التى كتبها الأطفال عن فكرتهم عن ذاتهم وتم تصنيفها إلى عدد من الفئات (١١ فئة) هى الخصائص الجسمية والمظهر - الملبس والهندام الصحة - المنزل والأسرة - الإستمتاع والترفيه - القدرة على أداء التمرينات الرياضية - القدرة المدرسية والإتجاه نحو المدرسة - القدرات العقلية - المواهب الخاصة (الموسيقى - الفنون) أن يظل كما هو - الشخصية والخلق والميول العاطفية

وصف المقياس:

يتكون المقياس فى صورته النهائية من ١٤٠ عبارة متضمنة ١٢ عبارة لقياس الكذب، وقد تم تطبيق المقياس على عينات (٦٠٠ من الإناث و٥٨٣ من الذكور) من الصفوف من الرابع الإبتدائى وحتى الصف الثالث الثانوى وقد تم حساب المتوسطات والإنحرافات المعيارية وقد تم إستبعاد عدد من الفقرات ليصبح عددها النهائى ٨٠

فقرة بعد إستبعاد الفقرات المكررة وفقرات قياس الكذب. وقد تم حساب الثبات باستخدام طريقة التجزئة النصفية وتراوحت معاملات الثبات بين ٠,٧٨-٠,٩٤ كما حسب معامل الثبات بإعادة التطبيق وبلغت معاملات الثبات ٠,٨٠-٠,٨٦ لعينات البنات ، ٠,٦٨-٠,٨٨ لعينة الذكور.

صدق الإختبار:

إستخدم فى تحديد الصدق التلازمى بعض المحكات مثل مقياس تقدير الذات على عينة من٣٤ مفحوصا وبلغ معامل الإرتباط ٠,٧٤ كما تم تحديد صدق المحتوى من خلال العرض على أربعة من المحكمين، وكان يمكن الوصول إلى مزيد من التحليل مع وجود عينات تقننين يمكن الحصول منها على بيانات أدق. كما تم حساب درجات مئينية لعينات البنات والبنين (ن = ١١٨٣) ومن الملاحظ أن جدول المعايير كان يحتاج إلى مزيد من التعمق السيكومترى من حيث عدم وجود فروق بين الصفوف الدراسية أو تمايز الجنس.

مقياس السلوك العدوانى للأطفال

إعداد آمال عبد السميع اباظة

يسعى مقياس السلوك العدواني للأطفال إلى قياس :

أ- العدوان المباشر : ويقصد به توقيع الأذى أو الضرر بالآخرين أو الذات ويتم التعبير عنه بطريقة مباشرة وواضحة وتشمل العدوان المادى Physical Aggression

ب- السلوك العدوانى اللفظى: ويقصد به الإستجابة اللفظية التى تحمل الإيذاء النفسى والإجتماعى.

ج- السلوك العدوانى غير المباشر: ويعبر عنه بطريقة إسقاطية على الذات أو على الآخرين.

ويتألف المقياس من ثلاثة أجزاء كل منها يقع فى ١٤ عبارة يتم الاستجابة عليها فى تدريج رباعى. وقد قامت معدة المقياس بتحديد صدقه بعرضه على مجموعة من الخبراء من جامعة طنطا، أما الثبات فتم حسابه بإعادة التطبيق على ٦٠ مفحوصا من المرحلة الإعدادية وتراوحت معاملات الثبات بين ٠,٧٥- ٠,٨٠ . كما ذكرت المعدة

أنهـا قـامت بـحـسـاب الصـدق الـتـمـيـيـزى بـحـسـاب دلالة الفـروق بين المتـوسطات لمجموعتين طرفيتين ولم يتم تحديد إجمالى عدد العينة المشتق منها العينات الطرفية كذلك لم يتم إختيار المجموعتين بناء على محك مستقل وعليه فما أشير عليه أنه صدق تمييزى لا يعتد به كما أوردت معدة المقياس معاملات الإرتباط بين الأنواع الثلاثة من العدوان ولم تتحدد علاقة كل نوع بالدرجة الكلية للمقياس . وقد تضمنت كراسة تعليمات الإختبار جدولا يضم درجات معيارية معدلة (لم يتم تحديد نوعها) كما تم فصل الذكور عن الإناث دون الإشارة الى ما إذا كان هناك فروق بين الجنسين من عدمه.

إختبار الإسلوب المعرفى

Cognitive Style Test (C.S.T)

أعدت هذا الإختبار آمال عبد السميع اباظة عن أصله الأجنبى الذى أعده كل من «بلاكـيـورن»، «جـونز» و«لـيـفين» بعـد أن قـام كل من «ويلكنسون و«بلاكـيـورن» بتطويره ويسعى الاختبار إلى قياس الإسلوب المعرفى للإكتئابين وتصنيف الأحداث تبعا لوجـهة نظر بيك فى نظريته المعرفية للإكتئاب الثلاثى فى (الذات – المستقبل – العالم) .

وهذه الإتجاهات الثلاثة مقسمة إلى نظرة سارة، ونظرة غير سارة، وكل محور فـرعى يشـمل خـمس بنود وكل إسـتـجـابة تـعكس درجـة الـتـحـريف أو الخلل المعـرفى للاكتئاب والأخطاء المعرفية المتضمنة فى النظرة غير السارة فى المحاور الثلاثة كما حددها «بيك» وهي:

– الإستنتاج الإستبدادى Arbitrary inference

– التجريد الإنتقائى Selective Abstraction

– المبالغة فى الجوانب السلبية Magnification

– التقليل من قيمة الأشياء Minimization

– إعتبار الملاحظة موجهة ضد الذات Personalization

والإختبار قنن على عينات انجليزية (م = ٧٠,٣، ع = ١٠,٨) والأسوياء (م =

٥٨,٦ ، ع = ٨,٤) والدرجة العالية تدل على السلبية فى الإسلوب المعرفى . ويتألف الإختبار من ٣٠ عبارة كل محور به ١٠ مفردات منها خمسة مفردات لقياس النظرة السارة (P) والخمسة الأخرى للنظرة غير السارة (U) للمحاور الثلاثة (الذات S) و(المستقبل F) و(العالم W) ويقوم المفحوص بالإجابة على المفردة بإختيار من متعدد على متصل رباعى. وقد قامت معدة الإختبار بحساب معاملات الإرتباط بين كل محور والدرجة الكلية للإختبار (ن = ٥٤ طالبة، ١٢٠ طالب) حيث تراوحت بين (٠,٧١ ، ٠,٨٩) للنظرة السارة (٠,٧٧ ، ٠,٨٩) للنظرة غير السارة فى المحاور الثلاثة، أما الثبات فقد تم حسابه بطريقة إعادة التطبيق على عينة قوامها ٣٥ طالبا وطالبة للدرجة الكلية ٠,٧٦ للطلاب، ٠,٨١ للطالبات.

مقياس التفكير الأخلاقى المصور للأطفال

إعداد فوقية عبد الفتاح

وهو مشتق عن اختبار الأحكام الأخلاقية Moral Judgement test والذى اعده بياجيه فى اطار نظريته عن النمو الأخلاقى عند الأطفال، وطوره رونالد جونسون بتطويره ١٩٦٢ وخضع المقياس لدراسات متعددة قام بها كولبرج. والمقياس يتكون من ٢٠ قصة تستثير اهتمامات الاطفال وعلى مستوى متكافئ فى مستوى صعوبتها، وقد قام بإعداده للبيئة المصرية طلعت منصور غبريال وحليم بشاى سنة ١٩٨٢ .

ويتألف المقياس الحالى من خمسة قصص مشتقة من المقياس الأصلى وتشمل مجالات الأحكام الاخلاقية الخمس كما حددها بياجيه، ويتم عرض المقياس فى تعديله الجديد مصورا لتيسير فهمه. والمقياس يصلح للاطفال من سن ٤-١٣ سنة حيث تعرض القصة ويعقبها ثلاثة اسئلة يتم الاجابة عنها بطريقة الاختيار من متعدد (٣ بدائل) وذلك بوضع علامة (صح) أمام الاجابة التى يختارها الطفل والمفردات الخمس هى :

أ- الجزاء فى مقابل العقاب التعويضى.

ب- المسئولية الجماعية

جـ- العدالة التلقائية

د- قوة تأثير العقاب

هـ- الواقعية الخلقية

ويتم تصحيح الاستجابات بإعطاء درجة واحدة للإستجابة الناضجة فى مقابل صفر للاستجابة غير الناضجة وبذلك تكون الدرجة القصوى للمقياس (١٠ درجات).

تقنين المقياس:

بعد عرض المقياس على محكمين للتأكد من صحته، وكذلك بعد عرضه على الأطفال فى الفئات العمرية المختلفة للتأكد من وضوح التعليمات وتحديد الزمن اللازم للإجابة سواء بصورة فردية أو جماعية، تم تطبيق المقياس على ٢٠٠ طفلا (٥٠ طفلا فى كل مجموعة عمرية ٤-٦، ٥-٧، ٨-١٠، ١١-١٣) ونصف هذه العينة من البنات تم اختيارهم من مرحلة رياض الأطفال والمرحلة الابتدائية والإعدادية من محافظة بنى سويف.

وقد تم حساب ثبات المقياس بطريقة اعادة التطبيق بفاصل زمنى ١٥ يوما وبلغ معامل الثبات ٠,٧٦ وهو معامل دال.

وقد تم تحديد صدق المقياس باستخدام التحليل العاملى حيث تشبعت بالعامل الأول جميع مفردات المقياس، والعامل الثانى تشعبت به جميع المفردات ماعدا المفردة الثالثة ولم يتضمن دليل المقياس اشارة إلى وجود معايير لتفسير الدرجات.

مقياس كامدكس Camdex (لقياس خرف الشيخوخة)

إعداد خالد فؤاد ابو حطب

The Camdex Examintion For Mental Disorder

يعد مقياس كامدكس Camdex مقياسا مقننا ومقابله شخصية مقننة لتشخيص الخلل العقلى فى الاعمار المتقدمة وبصفة خاصة لخرف الشيخوخة Dementia. ويتميز هذا الخلل عن باقى الأنماط الشائعة الأخرى من أمراض الشيخوخة وأيضا تمييز الحالة السوية ويمكن تطبيق هذا المقياس فى المجتمع أو المستشفيات ويمكن أن يقوم بتطبيقه الاطباء أو المتخصصون فى الصحة العقلية. أما مقياس «كامدكس -ر Camdex -R فهو صورة معدلة لهذا المقياس والذى شاع إستخدامه على مستوى

العالم، وكان شخيص خرف الشيخوخة معتمدا على بعض المعلومات المتاحة من المقابلات الشخصية المقننة وبعض الفحوصات ولا يقيس مقياس «كامدكس» «وكامدكس – ر» كل أنماط الخلل – الطبنفسى Psychiatric والذى يظهر فى المرحلة الثالثة من العمر (الشيخوخة) . ويعد مقياس كامدكس مقياسا شاملا لتشخيص الخرف، أما باقى الأعراض الجسمية أو الوظيفية والتى تتضمن القلق والمخاوف المرضية فيتم التمييز بينها وبين الخرف.

يختلف «كامدكس – ر «عن مقياس» «كامدكس» فيما يلى:

١– القدرة على تشخيص الخرف Dementia متضمنا نمط الخرف.

٢ – القدرة على التشخيص الاكلينيكى المعتمد على محكات التشخيص DSM-IV& ICD10

٣– يتضمن المقياس مجموعة من المفردات لقياس الذاكرة المجردة وتتميز هذه المفردات بالتماسك.

٤– الاستفادة من تقنين المقياس حول العالم وبصفة خاصة ما قامت به جاردن ابردين فى المملكة المتحدة من إجرائية القياس الوظيفى.

٥– يتضمن المقياس المعدل اسطوانة ممغنطة DIsk لتسهيل استخدام المادة والتصحيح وأيضا التحليل الخاص بنتائج التطبيق.

محتوى مقياس كامدكس – ر:

يتكون المقياس من :

١– مقابلة شخصية مقننة كلينيكية يتم فيها التعرف على المفحوص والحالة المرضية وتاريخها.

٢– مقياس للجوانب المعرفية ويتألف من مجموعة من الصور الجزء الأول منها يتطلب تعريفا والثانى يتطلب فهما لجمل مقروءة وتنفيذ الأمر المكتوب، والجزء الثالث استدلال منطقى حيث يختار المفحوص الاجابة من بدائل، والرابع التعرف على شخصيات من صورها والجزء الأخير فيه يتطلب تسمية أشياء مألوفة.

٣- تسجيل موضوعى للملاحظات الخاصة بالحالة العقلية للمفحوص وتسجل فى جدول معين.

٤- فحص جسمى وعصبى.

٥- فحوصات تحليلات معملية.

٦- الأدوية التى يتناولها المفحوص.

٧- بعض المعلومات الاضافية عن المفحوص.

٨- بعض المعلومات من الأقارب أو ذوى المفحوص.

٩- بعض المعلومات الأخرى إن وجدت.

زمن تطبيق المقياس:

يمكن تطبيق المقياس للجوانب التى تتعلق بالمفحوص فى ساعة واحدة ويمكن أن يقل الزمن فى الحالات غير المرضية، أما باقى المعلومات فتتطلب حوالى ٢٠ دقيقة.

تصحيح المقياس

يفضل أن يتم التصحيح متآنيا مع التطبيق وهناك بعض التعليمات متضمنة فى كراسة التعليمات عندما لا يستطيع المفحوص الإجابة عن سؤال ما حيث يقوم الفاحص بالإستعانة بالتعليمات المتضمنة على ال CD وتطبيق المقياس على الحالات المفحوصة تتطلب الحصول على موافقة من المفحوص للحالات المرضية المتوسطة والمتقدمة، ويمكن الحصول على الموافقة من الشخص المسئول عن رعاية المريضة أو المريض.

ثبات مقياس كامدكس - ر:

تم تطبيق مقياس كامدكس على عينات عمرية تمتد أعمارها بين ٧٧ سنة إلى ٩٠ + (ن = ٤١٨) هوبرت وآخرون (Huppert et al 1995)[١] ووجد أنه يتمتع بدرجة عالية من الإتساق ويمتد معامل الثبات بين (٠,٨٢-٠,٨٩) عندما طبق على عينات مختلفة، كما حسب معامل الثبات بإعادة التطبيق وتراوحت المعاملات بين (٠,٤٦-٠,٨٠) .

---

(١) Huppert, F.A. & etal, 1995, A Concise Neuropsychological test to assess Dementia. Dementia Diagnosis: socio - demographic deterinants in an elderly population sample. British J of clinical psychology, 34, 529-541

صدق المقياس:

أشار هوبرت ١٩٩٥ ان درجات المقياس (الدرجة الكلية وأيضا الدرجات الفرعية) تميز إلى درجة كبيرة بين ذوى خرف الشيخوخة وغيرهم من الأسوياء من ذوى الخرف البسيط أما المعايير فقد تم تضمينها فى صورة مستويات للتمييز بين الخرف Cut points. كما أشرنا ان المقياس استخدم عالميا، كما ان خالد فؤاد أبو حطب يقوم بالفعل بإستخدامه فى تشخيص حالات الخرف وجارى إعداد تقرير عن الحالات التى تم مناظرتها فى مصر باستخدام هذا المقياس الهام.

مقياس دافعية الإنجاز (صفاء الاعسر وآخرون):

– أعد هذا المقياس فى عام (١٩٨٣) وقد جرى استخدامه وتقنينه فى العديد من الدراسات التالية، مثل دراسة امنة عبد الله تركى (١٩٨٥) حيث قامت الباحثة باختصار المقياس إلى ٦٩ عبارة تندرج تحت ثمانية ابعاد رأت انها الأكثر أهمية بالنسبة لدافعية الإنجاز، ثم قامت بحساب الثبات فقط للمقياس بعد تعديله وذلك على عينة قوامها ٣٠ طالبة من طلاب المرحلة الثانوية.

– وقام نظام سبع النابلسى (١٩٨٦) بإعادة تقنين المقياس على عينة مكونة من ٣٠ طالب، وبناء على ذلك تم استبعاد ثمانية ابعاد نظرا لارتباطاتها السالبة أو المتدنية فإقتصر المقياس على عشرة أبعاد.

– وفى دراسة محمد توفيق عليوة (١٩٩١) قام الباحث باستخدام المقياس بعد حذف ٩ أبعاد، إلا انه لم يقم بإعادة تقنينه.

– كما قامت نجلاء عبد الله الكلية (١٩٩١) بتطبيق المقياس كاملا بعد حساب صدقه التلازمى مع مقياس (محمد اسماعيل عمران – ١٩٨٢) وذلك على عينة مكونة من ٣٠ طالبا، ٣٠ طالبة بالمرحلة الثانوية وكان معامل الارتباط بينهما قدره ٠,٦٧٤ وهو دال عند مستوى ٠,٠١ كما تم حساب معامل ثبات المقياس بطريقة اعادة التطبيق على نفس العينة السابقة، حيث كانت معاملات الارتباط للطلاب ٠,٥٠ والطالبات ٠,٤٩ وهى دالة عند مستوى ٠,٠١ .

– أما دراسة سحر محمد فتحى (١٩٩٤) فقد تم فيها استخدام المقياس بعد ان

قامت الباحثة باختصاره إلى ٥٠ عبارة فقط مع حذف بعدى القلق المعوق، والتقييد الوالدى، مع حساب الثبات فقط للصورة المختصرة بطريقة إعادة التطبيق على ٢٦ طالبة بكلية البنات، وبلغ معامل الإرتباط ٠,٧٤، بمستوى دلالة ٠,٠١ .

– وفى دراسة هشام ابراهيم (٢٠٠٣) قام الباحث بإعداد صورة مختصرة من المقياس مكونة من ٥٥ عبارة موزعة على عشرة أبعاد رأى الباحث انها الاكثر شيوعا واهمية بالنسبة لدافعية الانجاز وذلك بالرجوع الى التراث النفسى والدراسات السابقة التى تناولت ابعاد هذا المفهوم، وقد جرى تقنين المقياس بالتحقق من صدق استخدام الاتساق الداخلى بحساب معاملات الارتباط البينية لدرجات كل عبارة والدرجة الكلية للبعد الذى تنتمى إليه، وكذلك معاملات الارتباط البينية لدرجات كل بعد والدرجة الكلية للمقياس، وذلك على عينة مكونة من ٧٣ طالبا وطالبة بالمرحلة الثانوية الصناعية، كما تم التحقق من الثبات بطريقة إعادة التطبيق على عينة مكونة من ٤٠ طالبا وطالبة بنفس المرحلة، وكان معامل الثبات ٠,٨١٥ وهو دال عند مستوى ٠,٠١ .

الفصل الرابع عشر

قياس الشخصية عن طريق الأحكام والملاحظات المنظمة

رأينا فى قياس الشخصية عن طريق التقرير الذاتى أن الشخص يختار ويعطى تقديرات يرى أنها تعبر عن نفسه، أما عند قياس الشخصية عن طريق الأحكام Judg-ments والملاحظات المنظمة systematic observations فإن الذى يقوم بتقدير السلوك أوالأداء المميز للفرد أو المفحوص شخص أو أشخاص آخرون، ويمكن أن نقسم هذا النوع من قياس الشخصية إلى قسمين:

(أ) مقاييس التقدير وتقريرات القياس الاجتماعى

Ratings and Sociometric Reports

(ب) الملاحظة المنظمة لعينات السلوك: Systematic Observation of Be-havior Samples

(جـ) الاختبارات الموقفية ولعب الادوار Sisuational Tests and Role laying

وسوف نتناول كلا من هذين القسمين بشىء من التفصيل.

(أ) مقاييس التقدير وتقريرات القياس الاجتماعى.

يمكن أن نصنف مقاييس التقدير وتقريرات القياس الاجتماعى إلى نوعين على أساس من يقومون بهما، النوع الأول يقوم به الرؤساء أو الكبار، بينما يقوم الأفراد أو الزملاء بالتقدير والحكم فى النوع الثانى.

أما تقديرات الرؤساء أو الكبار فنحن نلاحظ أن الرؤساء أو المدرسين مثلا يقومون باستمرار بتقديم نوع من الأوصاف لمرءوسيهم أو لمن هم فى كفالتهم، إلا أننا لا نستطيع أن نقارن هذه الأوصاف لأنها تختلف فى أسلوبها، لهذا استعملت مقاييس التقدير حتى يتوافر قدر معقول من التشابه بين التقديرات.

ويتكون مقياس التقدير من قائمة من السمات يجرى عليها التقدير. وقد يتكون المقياس من قائمة بسيطة من الصفات يتم الاختيار من بينها، أو قد يكون مقياسا متصلا يتضمن أقساما وصفية متعددة للاختيار منها ، ويوضح المثال الثانى، (فى ٢٧ : ٥٠٧) هذا النوع من مقاييس التقدير المتصلة:

	* يسعى إليه الآخرون * يحبه الآخرون * يحتمله الآخرون	
١ – كيف تتأثر أنت والآخرون بمظهره وسلوكه؟	* يتجنبه الآخرون * يسعى إلى إضافة أعمال جديدة إلى نفسه * يكمل العمل الإضافى المقترح	وقائع مؤيدة
٢ – هل يحتاج إلى تحريك أو أنه يعمل من تلقاء نفسه؟	* يعمل واجباته العادية تلقائيا * يحتاج إلى تحريك أحيانا * يحتاج إلى تحريك كثير ليعمل واجباته العادية * لامجال لملاحظته	وقائع مؤيدة

ولكن ـ نلاحظ أن مقاييس التقدير عن طريق الرؤساء تكمن فيها مصادر متعددة للخطأ منها مثلا ما يسميه كرونباك (٢٧) «خطأ الكرم» ويقصد به ميل من يقوم بالتقدير إلى إعطاء تقديرات طيبة مؤيدة، فإذا سئل مدرس مثلا أن يقدر تلميذا ما فيما يختص بالميل إلى التعاون، فإنه سوف يقدر كل التلاميذ الذين لا يثيرون الشغب تقديرا عاليا جداً، فى أقصى طرف المقياس بينما يهبط بتقدير القلة من التلاميذ مثيرى الشغب، ويكون الارتفاع فى تقدير الغالبية رد فعل لتمييزهم عن مثيرى الشغب. ومن أسباب خطأ الكرم أيضاً أن الرئيس المقدر قد يحس أن مما يمس قدرته الرئاسية أن يعترف أن مرؤسيه ليسوا أكفاء ، أو أن الرئيس قد يجد من الأيسر عليه أن يحكم أحكاما طيبة على مرؤسيه إجمالا بدل أن يتملى ويطيل النظر ليميز بينهم تمييزا دقيقاً.

ومن مصادر الخطأ أيضاً فى هذا النوع من التقدير الغموض والإبهام، فقد يفسر المقدر سؤالا أو مصطلحا تفسيراً مختلفا، فقد يفسر شخص «القيادة»، مثلا على أنها تركيز السلطة فى يد القائد، والانفراد باتخاذ القرارات، فعندما يقدر هذا الشخص

مفحوصاً تتوافر فيه هذه الصفات فإنه يقدره تقديرا عاليا ومخالفا لتقدير شخص آخر للمفحوص ذاته، لأن المقدر الأخير يفهم القيادة على أنها تشجيع للمرءوسين ومساعدة فى اتخاذ قرارات جماعية.

ويعتبر أسلوب الاستجابة Response style عند المقدر مصدراً آخر من مصادر الخطأ فى مقاييس التقدير بواسطة الرؤساء أو الكبار، فمثلا نجد أن مقدرا ما يندر أن يستعمل أطراف المقياس فى وصف المفحوصين، بينما يميل شخص آخر إلى الإسراف فى استعمال هذه الأطراف، أى يميل إلى تصنيف المفحوصين إلى أقسام متميزة جداً، ومن ثم نجد أن مثل هذين المقدرين يختلفان عندما يقدران شخصاً واحدا.

وما يؤثر فى تقديرات الرؤساء أيضًا معلوماتهم عن المفحوصين قد تكون محدودة، فقد يختلف تقدير مدرس التربية الفنية ومدرس اللغة العربية فى تقدير ردود فعل تلميذ معين للهبوط، ويرجع الاختلاف إلى أن كلا منهما يعرف التلميذ ويلاحظه فى عدد محدود من المواقف، وعلى هذا تكون معرفته به محدودة وليست شاملة.

ومن مصادر الخطأ فى هذا النوع من التقدير «أثر الهالة» أو «وقع الهالة halo effect» ويحدث هذا الأثر أو الوقع عندما يتأثر المقدر تأثيرا قويا بالانطباع العام عن المفحوص، فيتكون عنده رأى عام أو تقدير عام عن قيمة الشخص يؤثر كثيرا فى تقديره لسمات هذا الشخص الخاصة، وهذا النوع من الخطأ يجعل نمط السمات فى الفرد غامضاً عند المقدر.

إلا أنه من الممكن تلافى كثير من مصادر الخطأ هذه إذا راعينا:

١ – الاختيار السليم لمن يقومون بالتقدير، من حيث قدرتهم ودرجة اندماجهم وتأثرهم شخصيا بتقدير للمفحوصين، فقد سبقت الإشارة إلى أن المقدر قد يميل فى تقديره إلى التحيز لصالح مرؤوسيه لأن فى هذا تأييد لكفاءة رئاسته.

٢ – كلما زادت معرفة المقدر بالمفحوصين كانت تقديراته أكثر صدقا.

٣ – إذا استعملنا مقدراً بعينه أكثر من مرة فمن المفيد أن نحتفظ بتقديراته لمقارنتها والتعرف على أسلوب استجابته، أو ما يمكن أن نسميه خطأه الثابت.

٤ – من الممكن زيادة ثبات مقاييس التقدير عن طريق الاعتماد على أكثر من مقدر واحد، على أن يكونوا على معرفة كافية بالمفحوصين.

٥ – مما يزيد من قيمة مقاييس التقدير أن تعد بعناية ودقة. وبصفة عامة نجد أن المقياس ذا الخمس نقط يعطى تمييزا أكثر دقة من المقياس أو القائمةالتى تعتمد على إجابات «نعم» أو «لا». كما أن مقياس الخمس نقط ينبه إلى أنواع متباينة من الاختلاف والانحرافات، وكذلك يساعد الاختيار الإجبارى الذى تكلمنا عنه فى بعض اختبارات التقدير الذاتى، على زيادة قيمة مقياس التقدير، والهدف من الاختيار الإجبارى هو فصل عمل المقدر فيما يتصل بوصف ما يعمله المفحوص عن تقويم هذا العمل، أى أن المقدر يكون مسئولا عن وصف المفحوص، بينما يقوم المسئول عن اتخاذ القرار بتقويم هذا النوع من مقاييس التقدير القائم على الاختيار الإجبارى أكثر ملاءمة لقرارات المؤسسات فيما يتصل بالانتقاء والتصنيف، ولكنها لا تناسب كثيرا عملية التوجيه التربوى أو الوصف التفصيلى للفرد.

أما عن صدق مقاييس تقدير الرؤساء فتشير الدراسات التى أجريت عنها إلى ضرورة أخذ هذه المقاييس بحذر فيما يتصل بالحكم على الأفراد من جهة، وباستعمالها للتحقق من صدق بعض مقاييس واختبارات الشخصية من جهة أخرى، ومن أمثلة هذا أن طريقة تقدير الكبار استعملت فى دراسة صدق مقياس المسئولية الاجتماعية، الصورة ت، الذى تناوله الفصل السابق، حيث قام المدرسون بتقدير تلاميذ العينة على مقياس خاص بالمدرسين وكان معامل الارتباط بين تقدير المدرسين ودرجات التلاميذ على المقياس ٠,٤٦ وقد اعتبر هذا المعامل مقبولا ،وخاصة إذا أخذنا فى الاعتبار أن تقدير المدرس للتلميذ فى أى ناحية من نواحى السلوك تتدخل فيه عوامل كثيرة منها على سبيل المثال، تحصيل التلميذ الدراسى، ومستواه الاقتصادى الاجتماعى، ودرجة جاذبية التلميذ الشخصية، وتفضيل المدرس وتحيزه الشخصى وهكذا، (٢ : ٢٧).

وقد أثبتت دراسة قام بها بيترز وكاميل (٥٢) أن ارتباط تقديرات الرؤساء لمدى معرفة العاملين ودرايتهم بأعمالهم مع هذه المعرفة والدراية كما قيست باختبار خاص، كان الارتباط بينهما ٠,٣٥ بينما وصل هذا الارتباط فى بعض الأقسام ٠,٥٥ كما أظهرت دراسات أخرى ارتباط التقديرات بطول المدة التى يعرف فيها المقدر المفحوصين، وبمقدار حبه لهم (فى ٢٧ : ٥١٨).

ولكن على الرغم من حدودمقياس تقدير الرؤساء والكبار، إلا انها تعد مصدر غنيا بالمعلومات لا يمكن الحصول عليها إلا عن طريقها. ومن قوائم الملاحظة الحديثة ما قامت مها فؤاد أبو حطب بإعداده حيث قامت باعداد وتقنين قائمة ملاحظة سلوك الرفض المدرسى والأعراض المرتبطة به **School Refusal Behavior Symptom Check list (SRBS -C)** حيث يعتبر رفض المدرسة أحد أعظم المشكلات الإنفعالية التى قد تواجه الأطفال والمراهقين والتى عادة ما يكون لها العديد من الآثار السلبية على الطفل والأسرة وفريق العمل فى المدرسة كما يرتبط رفض المدرسة فى كثير من الأحوال بالعديد من الاضطرابات النفسية التى عادة ما يكون هذا السلوك أحد مظاهرها.

وقد هدفت الدراسة التى قامت بها معدة القائمة (١٧-أ) وموضوعها «رفض المدرسة: التقييم لدى عينة مجتمعية»(١٧- جـ) الى رصد هذه الظاهرة وتحديد الاعراض والإضطرابات المرتبطة بها.

وقد تم اختيار (١٥٠ أم) والتى تم تقييم أبنائهن الذين تقل أعمارهم عن ١٨ سنة والذى بلغ عددهم ٢٦٦ مفحوصا، وقد تم هذا التقييم من خلال قائمة الملاحظة التى اعدتها الباحثة لأغراض هذه الدراسة.

وتتضمن هذه البطاقة من الاعراض الشائعة والسلوك غير المقبول والذى عادة ما يظهر فى المنزل والمدرسة ويرتبط بسلوك رفض المدرسة، وهذه القائمة تطبق فرديا وتتضمن مقياسين فرعيين هما:

أ- **مقياس سلوك الرفض المدرسى** School Related Refusal Behaviour scale RRB-s

ويتضمن اعراض الرفض المدرسى والتى تتألف من اعراض السلوك المضطرب والمستمدة من المعايير التشخيصية للدليل التشخيصى الإحصائى الرابع DSM وهى:

أ- الاداء فى المدرسة.

ب- مشكلات التعلم.

ج- علاقات الاقران

د- علاقة الطفل/ المراهق بالمعلمين

٢- مقياس الاضطرابات السلوكية المرتبطة بالرفض المدرسى Disorder Related Refusal Behaviour Scale (DRRB-S) ويتضمن أعراض الرفض المدرسى، وتعتمد على اعراض السلوك المضطرب المستمدة من DSM وهى:-

أ- اضطراب القلق المعمم.

ب- المخاوف المرضية من المدرسة

ج- المخاوف المرضية من المجتمع

د- الإكتئاب

كما استخدمت الباحثة المحكات التشخيصية للأعراض السابقة كما تتحدد بمؤشرات DSM (١٧-ب)،

وقد تم تقنين قائمة ملاحظة سلوك الرفض المدرسى والاعراض المرتبطة على عينة قوامها ١٢٨ طفلة، ١٣٨ طفلا (ن = ٢٦٦) تتراوح أعمارهم بين ٤-١٥ سنة وتم تحديد صدق القائمة بحساب معامل الارتباط بين الدرجة التى قدرت للمفحوص فى مقياس سلوك الرفض ومقياس الاضطرابات السلوكية والمحك DSM ٠,٥١، ٠,٨٦٥ وعلى التوالى ، أما الثبات فقد حدد بملاحظة الاستجابات المتسقة للمفحوص والتى تكررت فى مناسبات مختلفة.

تقنين قائمة كونر لتقدير المشكلات السلوكية للأطفال Conner's Rating Scale

فى دراسة قامت بها مها فؤاد أبو حطب (١٧-أ) هدفت الى كشف المشكلات السلوكية للأطفال فى مرحلة ما قبل المدرسة، وبصفة خاصة أولئك الاطفال الذين يعانون من مشكلات نقص الانتباه المصحوبة أو غير المصحوبة بفرط الحركة لدى عينة عشوائية من أطفال ما قبل المدرسة من مدينة القاهرة (من ٨ مدارس مختلفة من منطقتى المعادى وحلوان) وقد تم تقدير سلوك الاطفال من خلال معلميهم (ن =٢١) وتم اختيار ١٠٢ طفلا، ٩٦ طفلة والذين يتراوح عمرهم بين ٤-٦ سنوات.

وقد استخدمت الباحثة قائمة كونر لتقدير المشكلات السلوكية للأطفال - نموذج المعلم وهذه القائمة أعدها السمادونى ١٩٩٥ وتتألف من ٣٧ مفردة تحصر المشكلات

السلوكية الشائعة للاطفال من سن ٤-١٤ سنة، وتستخدم فى التشخيص والعلاج.

وقد تم استخلاص ستة مقاييس فرعية من القائمة الاساسية لقياس فرط الحركة - القلق وسوء التوافق الاجتماعى، نقص الانتباه، الإكتئاب- الطاعة.

وقد تبين وجود أعراض كل من نقص الانتباه بنسبة ٦,١ ٪ وسوء التوافق الاجتماعى بنسبة ٥,٦ ٪ ، فرط الحركة ٤,١ ٪ الاكتئاب ٣,٦ ٪ والقلق ٣ ٪ بدرجة شديدة لدى عينة الدراسة، ووجدت فروق بين الذكور والاناث وبصفة خاصة فرط الحركة ونقص الانتباه والقلق ، بينما كانت الاناث أكثر اكتئابا وأكثر عرضة لاعراض سوء التوافق الانفعالى.

وقد حددت الباحثة الصدق العاملى للقائمة وحصلت على ستة عوامل هى : النشاط الزائد، القلق، ضعف التكيف الاجتماعى، نقص الانتباه، الاكتئاب وضعف التكيف الانفعالى والاندفاعية. كما اعتمدت على تحديد اشتراكيات العوامل كمؤشر للثبات.

٢ - مقاييس تقدير الأقران:

والنوع الثانى من مقاييس التقدير هو تقدير الأقران peer ratings التى تعطى معلومات تعلو فى قيمتها على تقديرات الكبار أو الرؤساء، وحتى عندما تتوافر تقديرات الرؤساء فإن تقديرات الأقران والزملاء تعطى بيانات عن جانب مختلف من الشخصية، فإذا كان رئيس واحد هو الذى يقوم بالتقدير فإن قيام مجموعة من الأفراد أو الزملاء بالتقدير سوف يعطى فرصة لعدد كبير من الأحكام مما يترتب عليه أن متوسط التقديرات بالنسبة إلى سمة معينة من سمات الشخصية سوف تكون أكثر ثباتا من التقديرات الذى يقوم به رئيس واحد، ومن الممكن أن ننظر إلى تقديرات الأفراد على أنها تعبير موضوعى عن سمعة الفرد، فالسمعة تقوم إلى حد ما على السلوك الفعلى، إلا أن نمط العلاقات الاجتماعية فى داخل الجماعة لا شك يؤدى إلى وجود نوع من التحيز.

ومن الأساليب المستعملة فى تقدير الأفراد ما يسمى أسلوب تعيين الأسماء nomination أو الترشيح، ويستعمل هذا الأسلوب كثيرا فى القياس الاجتماعى Sociometry. وفى هذا الأسلوب يسأل كل عضو من أعضاء الجماعة أن يسمى عددا محددا

من الأشخاص الذين يتميزون بشكل واضح فى ناحية من النواحى، مثل القيادة، كما يسمى أيضاً الأشخاص الذين تنقصهم هذه الخاصية، وبالنسبة إلى الأطفال فإن عملية التسمية قد تتخذ صورة لعبة التخمين، أو اختبار «خمن من» Guess Who الذى يصف الأدوار المختلفة التى يمكن أن يلعبها الأطفال ويسمى كل عضو فى الجماعة أسماء الأطفال الذين يعتقدون أن كل وصف يلائمهم، وترسم لكل طفل صفحة نفسية عن طريق عدد المرات التى أشير إليه فيها لكل وصف.

وإذا كان من الممكن دراسة خصائص الفرد بطريقة «خمن من»، فإن الرسم الاجتماعى (السوسيوجرام Sociogram) يعطى فهما واستبصارا أعمق فى تمييز التجمعات الداخلية فى الجماعة، وبيان البناء الهرمى للقيادة والمنعزلين، أى هذا الرسم الاجتماعى يعتبر طريقة لدراسة البيئة الاجتماعية فى الجماعة (١٦: ٤٤٣ وما بعدها).

والذى أنشأ الرسم الاجتماعى وطوره مارينو Mareno (٢٩). إلا أن الأسلوب قد عدل بعدة طرق قللت من فاعليته فى سبيل السهولة واليسر فى الاستعمال، وأفضل طريقة فى استعماله هى أن يسأل أعضاء الجماعة أن يختاروا رفقاء لهم فى نشاط معين، ولابد أن تكون التعليمات متصلة بأنشطة حقيقية للجماعة، كما يجب أن تكون الاختبارات حقيقية ويجب أن يعلم أعضاء الجماعة أن المعلومات التى يعطونها سوف تصان سريتها، وأن البيانات المعطاة سوف تستخدم فيما أخبر أعضاء الجماعة أنها سوف تستخدم فيه، مثلا لاختيار لجان، أو جماعات عمل فرعية، أو تنظيم الجلوس فى فصل دراسى وهكذا.

ويعتمد الرسم الاجتماعى الذى نحصل عليه على السؤال الذى توجهه، فإن نمط الرسم الذى ينتج عن سؤال مجموعة من التلاميذ عن اختياراتهم لأقرانهم الذين يحبون أن يستذكروا معهم سوف يختلف عن نمط الرسم الناتج عن سؤال عن اختيارات من يحبون اصطحابهم فى رحلة.

ويشمل مصطلح التقدير بالقياس الاجتماعى Sociometric rating كل طرق تمييز العلاقات الاجتماعية بين أعضاء الجماعة. وليس هناك فصل قاطع بين تقدير الأفراد الوصفى، وبين التقدير بالقياس الاجتماعى، إلا أننا نجد النوع الأخير بصفة محدودة عمن يحبهم المقدر أكثر أو يفضل العمل معهم، ولهذا فهى تتصل باستجابة المقدر كما تتصل بشخصيته.

ولتقديرات الأقران استعمالات متعددة منها، أن قائد الجماعة يستعملها فى التعرف على الأشخاص الذين يحتاجون إلى عناية خاصة، أو يستعملها فى إعادة تنظيم الجماعة حتى تؤدى وظيفتها بشكل أفضل. كما أن هذه التقديرات يمكن أن تتخذ أساسا للانتقاء والتصنيف، وهناك من يعتبرها أنقى مقياس للقيادة، أما بالنسبة إلى المرشد النفسى. فإن تقديرات الأقران توضح خصائص الشخص التى تعوق قبول الجماعة له، وخاصة عندما ينقص الشخص الاستبصار الكافى فى رأى الآخرين فيه.

ويذهب كثير من الباحثين إلى أن تقديرات الأقران أثبتت أنها إحدى وسائل التقدير التى يمكن أن يوثق بها إلى درجة بعيدة لعدة أسباب هى:

١ – أن عدد المقدرين كبير، وقد أشرنا منذ قليل إلى أهمية هذا فى زيادة ثبات التقديرات بصفة عامة.

٢ – أن أقران الشخص عادة ما يكونون أكثر قدرة من غيرهم على ملاحظة السلوك المميز له، وهم من ثم أفضل فى الحكم على سمات معينة تتصل بالعلاقات الاجتماعية من المدرسين أو المشرفين أو أى ملاحظين خارجيين.

٣ – أن آراء أعضاء الجماعة الآخرين، صوابا كانت أم خطأ، تؤثر فى سلوكهم وتصرفاتهم وبالتالى فإن هذه الآراء تحدد جزئيا طبيعة تفاعلات الشخص مع الجماعة.. فمثلا، إذا كان أعضاء جماعة يرون أن أحد أعضاء جماعتهم، وليكن اسمه «علياء مثلا، رجل يمكن الاعتماد عليه والثقة به، فإن هذا الرأى يوجه ويحدد نوع تعاملهم مع «على»، وغالبا ما يحدد هذا العامل نوع السلوك من جهة «على» كما يحدد نوع التفاعل من جهة الجماعة، فالجماعة تعاملة على أنه ثقة، وهو يسلك تجاه الجماعة وفق توقعاته منهم (٥ : ٩٨ – ١١٠) ولاشك أن هذا يعطى تقديرات الأقران صفة أقرب إلى الواقعية.

(ب) الملاحظة المنظمة لعينات السلوك

من الممكن أن تعطينا الملاحظات المنظمة وصفا أكثر دقة للسلوك أو الأداء المميز إذا قورنت بأساليب التقريرالذاتى أو الأحكام والتقديرات التى يصدرها الآخرون عن الشخص، ذلك لأنه فى التقرير الذاتى من المحتمل أن يتدخل التذكر الانتقائى، بل أنه يحدث أيضاً عندما تستعمل أساليب التقدير والحكم المختلفة، وكذلك نجد أن الملاحظة

فى هذه الأساليب الأخيرة ليست قائمة على الاختيار المنظم للمواقف التى يلاحظ فيها السلوك، بل أنها تقوم على ملاحظة المواقف كيفما اتفق وقوعها، وفى هذه الأساليب أيضًا لايلاحظ الشخص فى كل المواقف، لهذا قلنا أن الملاحظة المنظمة تعطى وصفا أدق للسلوك المميز.

ومن السهل أن نلاحظ سلوك مفحوص ما فى ظروفه العادية، وهذا ما يسمى بالملاحظة الميدانية، هذا النوع من الملاحظة يكون أكثر ملاءمة فى بعض الدراسات من الملاحظات المقننة، فكثير من الباحثين يرى أنه من المستحيل أن نفهم الشخصية إلا إذا لاحظنا المفحوص وهو يستجيب إلى الظروف التى لها مغزى بالنسبة إليه، لأن المثيرات المختلفة لها معان ومغاز مختلفة بالنسبة إلى الأفراد المختلفين، وعلى هذا فهم يرون أن المواقف المقننة حيث تلاحظ ردود أفعال الأشخاص لموقف بعينه، تبرز أنماط السلوك المهمة كما تبرزها الظروف العادية، غير المتشابهة، التى يعيش فيها الفرد، إلا أن المشكلة هنا تكمن فى كيف نرى قدرا كافيا من سلوك الفرد، والحصول على تسجيلات لملاحظات السلوك يمكن الاعتماد عليها. أى أن المشكلة هنا ذات شقين: الأول، عينة السلوك وما يرتبط بها من صعوبات، والثانى، عيوب أو أخطاء من يقوم بالتسجيل، ولنتكلم عن كل منهما بشىء من التفصيل.

اختيار عينة السلوك:

لكى تعرف السلوك أو الأداء المميز لشخص ما لابد أن نعرف ما يميز تصرفه فى موقف بعينه، ولكن المواقف تتغير من يوم إلى يوم ومن لحظة إلى لحظة أخرى، والطريقة الوحيدة لنطمئن إلى أن ما نلاحظه هو السلوك المميز هى أن ندرس هذا الشخص فى مواقف ومناسبات عديدة، وهو أمر مجهد ومكلف، فالواقع أنه لابد أن نصل إلى نوع من التوفيق بين الإختيار الدقيق لعينة السلوك والاقتصاد فى الإنفاق والجهد.

وأنه من العسير جدا أن نقوم باستنتاجات عن فروق بين الأفراد من ملاحظة سلوكهم فى مواقف معينة، لأنه لايمكن ملاحظة فردين فى موقف واحد بعينه، وحتى إذا كان الموقف واحد فى ظاهره فإن الظروف السابقة الخاصة بكل منهما تجعلهما يسلكان سلوكا مختلفا، ولهذا فإن أقصى ما يمكن أن تقدمه لنا الملاحظات

المقارنة هو بيان كيف يسلك أشخاص مختلفون فى ظروفهم الحالية، ولكن هذه الملاحظات لا تضمن لنا أن تستمر هذه الفروق عندما تتغير طبيعة هذه الظروف.

ومن أفضل الأساليب التى تؤدى إلى المقارنة الدقيقة ما يسمى بالعينات الموقوتة time sampling وفى هذا الأسلوب يخطط مسبقا جدول أو قائمة للملاحظات ويصمم هذا الجدول بشكل عشوائى يضمن أن يلاحظ كل مفحوص فى ظروف متشابهة مع الاخرين بحيث تمكن المقارنة. فمثلا عند دراسة الاتصالات الاجتماعية لأطفال ما قبل المدرسة يوضع جدول الملاحظات لمدة دقيقة واحدة، وبعد ملاحظة الطفل لمدة دقيقة يكتب الملاحظ تقريرا كاملا عما لاحظه من تفاعلات الطفل الاجتماعية، ويلاحظ الأطفال وفق نظام مرسوم ويعدل من يوم إلى يوم، وفى خلال الدراسة كان كل طفل يلاحظ عدد مرات متساوية فى خلال الخمس دقائق الأولى من ساعات اللعب الحر، ثم الخمس دقائق الثانية وهكذا. وللعينات الموقوتة قصيرة الأجل والموزعة توزيعا حسنا عدة مزايا ، فالصورة المتراكمة من الملاحظات فيها تكون أكثر تمثيلا للسلوك من نفس القدر من المعلومات التى يحصل عليها من ملاحظات من مرات أقل وذات فترات طويلة، كما أن أخطاء الذاكرة تقل هنا لأن الملاحظ يقوم بالتسجيل أثناء وبعد الملاحظة مباشرة، والعينات الموقوتة أكثر مناسبة لتسجيل حقائق معينة يمكن أن تصاغ فى صورة رقمية، مثل عدد مرات الاتصال الاجتماعى بالآخرين.

ومن الممكن الحصول على عينة أكثر اتساعا بأسلوب «السجل اليومى» day record الذى أنشأه باركر وزملاؤه (٢١)، كان هدفهم هو دراسة نمط حياة الطفل ومع الاهتمام الخاص بالمجالات التى يتحرك فيها.

وللسجل اليومى مزايا أن يعطى صور عن تتابع النشاط ، ولكن له عيوبه، فإذا كان الملاحظ معروفا ومكشوفا تأثر السلوك الذى يلاحظه، والاعتماد على يوم كامل بعينه لايكفى للحصول على معلومات مميزة كاملة عن المفحوص.

ومن المهم أن نلاحظ بالنسبة إلى عينة السلوك الملاحظة عن طريق العينة الموقوتة ما يلى:

١ – أن الاستجابات التى يعدها الملاحظ متشابهة أو واحدة قد يكون لها فى

الحقيقة معان مختلفة، وأن المواقف التى تبدو له متشابهة قد تستدعى استجابات متباينة.

٢ – أن الاستنتاجات المشتقة من موقف واحد، حتى ولو كانت على أساس تراكم ملاحظات، تكون أى الاستنتاجات صحيحة بالنسبة إلى هذا الموقف فقط.

٣ – أن الملاحظات المطلوبة والمناسبة للحصول على معلومات ثابتة ويمكن الثقة بها، يعتمد هذا العدد على المشكلة، إلا أنه بصفة عامة نجد أن عددا كبيرا من الملاحظات القصيرة أفضل من عدد قليل من الملاحظات التى تعتمد على فترات طويلة.

عيوب من يقوم بالملاحظة:

الواقع أن بعض عيوب الملاحظين قد يرجع إلى طبيعة المواقف الاجتماعية ذاتها التى تكون موضوع ملاحظة حيث تحول درجة تعقدها الكبيرة دون التقرير الشامل للمستغرق لها، أما العيوب الأخرى فترجع إلى الملاحظ نفسه.

وإذا كانت أخطاء الملاحظ مجرد حذف أو إغفال عشوائى لبعض ما يلاحظ فليس لهذا وزن كبير، وأما ما يجب الاهتمام به فهو الأخطاء المنظمة كأن يبالغ باستمرار وبصفة مطردة فى تأكيد بعض أنماط الأحداث، كما يغفل الإشارة إلى بعضها الآخر.

عندما يلاحظ مراقبون موقفا واحدا فإنهم يعطون تقريرات متباينة بدرجة كبيرة، ذلك لأن كل ملاحظ عنده حساسية لأنماط معينة من السلوك أكثر من غيرها والوضع الأمثل هو أن يبنى الملاحظ انطباعه على كل تصرف يمكن أن يكشف أو يعكس خصائص أساسية، ولكن عندما يركز من يقوم بالملاحظة والمراقبة على شىء بعينه فإنه بالضرورة لابد أن يغفل ملاحظة أشياء أخرى.

وثمة عيب آخر وهو أن الملاحظين يفسرون ما يرون، فإذا قام الملاحظون بتسجيل الحقائق الموضوعية، فقد يصل آخرون إلى تفسيرات مختلفة تماما، ولكن الناس يميلون إلى تفسير ما يرون، وعندما يفسرون فإنهم يميلون إلى تجاهل الحقائق التى لا تتفق مع تفسيرهم بل أنهم قد يخترعون حقائق لتكمل تفسيرهم.

ومما يقلل من آثار هذه العيوب أن يتم التسجيل بشكل منظم، وأن يكون كلما أمكن ، فى صورة وحدات سلوكية يمكن عدها، فمثلا من الممكن أن نصف مدى انتظام عمال مصنع فى عملهم عن طريق التسجيل الزمنى للحظات التى يكونون فيها فى عمل فعلى، واللحظات التى ينشغلون فيها عن العمل للحصول على الأدوات، أوزيارات زملائهم، أو يتعرضون فيها لمؤثرات التشتيت، وحتى إذا كان السلوك الملاحظ على درجة كبيرة من التنوع بحيث يصعب وضعه فى صورة وحدات سلوكية فمن الممكن أن تحدد أقسام السلوك وتعد قبل ويكون عمل الملاحظ هو أن يصنف السلوك الذى يلاحظه وفق هذا التقسيم، ويضع علامة فى القسم المناسب لكل أسلوب يلاحظه كما هو الحال فى أقسام بيلز Bales (راجع ٦).

وعلى الرغم من فائدة التسجيل والعد الموضوعى لوحدات السلوك وملائمته للبحوث العلمية فإن فائدته محدودة فى التوجيه والإرشاد النفسى.

أما السجلات القصصية أو الوقائعية anecdotal records فإنها وإن كانت تنقصها دقة الموضوعية، فإنها تعطى وصفا حيا للشخص، لأن الملاحظ فيها يوجه اهتمامه إلى أى سلوك يبدو جدير بالملاحظة بدل أن يركز على سمات معينة للمفحوصين كلهم إلا أن الملاحظ فى السجل القصصى يصف بدقة ما يلاحظه فاصلا بين الوقائع وتفسيراتها، وعندما تتراكم هذه السجلات القصصية فإنها تعطى صورة أكثر غنى من أى وسيلة أخرى على هذه الدرجة من البساطة، ومسئولية الملاحظ الكبرى هنا تتركز فى اختيار الوقائع التى تستحق التقدير. وفى أن يتحرى الموضوعية بقدر الإمكان.

(جـ) الاختبارات الموقفية ولعب الأدوار: Situational Tests and Role Playing

نشأ هذا النوع من الاختبارات فى مركز تخطيط الخدمات الاستراتيجية بالولايات المتحدة الأمريكية Unitded States office of Strategic services فى الحرب العالمية الثانية، وتتميز هذه الاختبارات بأنها تتبنى تقنية خاصة فى قياس وتقدير الشخصية وتقدير السمات المقاسة، فهى تتطلب معايشة خاصة للمفحوص لعدة ايام لمواقف معينة يتم ضبطها وتقنينها بحيث تتشابه مع المواقف الحياتية الطبيعية من الوجهة الفيزيقية والاجتماعية ويتم تقويم اداء المفحوص من خلال الملاحظة.

وقد استخدمت هذه التقنية فى انتقاء الجيش الامريكى الذين سيقومون بالخدمة العسكرية فى أماكن غير الولايات المتحدة. وقد استخدمت نفس التقنيات فى معهد البحوث والقياس للشخصية فى جامعة كاليفورنيا وايضا فى بعض المشروعات فى الانتقاء سواء للاغراض العسكرية أو المدنية أو فى مجال الصناعة وبصفة خاصة عند قياس المواقف الضاغطة Situational Stress test SST أو المواقف الحوارية للقادة Leader's Group Discussion LGD ويتطلب هذا النوع من الاختبارات الموقفية تدريبا جيدا للمقدرين الذين يقومون بملاحظة هذه المواقف، كما أن هذه المواقف إذا احسن تقنينها تصبح اداة جيدة للتنبؤ بالاداء فى المواقف التى تتطلب حوارا لفظيا، أو حل المشكلات اللفظية، أو لعب الادوار والتى تتطلب ارتجالا (أى بدون تحضير) للموقف الموجه والمعد لغرض معين، ويتم تقدير الاختبارات الموقفية واختبارات لعب الدور من خلال الملاحظة أو التسجيل بالفيديو.

ويوجد عدد قليل نسبيا من هذه الاختبارات وفيما يلى وصفا لبعض هذه الاختبارات.

اختبار حالة وسمة القلق للأطفال Trait state Anxiety Inventory for children اعداد عبد الرقيب أحمد البحيرى

هذا الاختبار مقتبس من اختبار Charles D.Spielbergar والمعروف باسم State - Trait Anxiety Inventory for children وذلك لقياس مفهومين منفصلين للقلق أطلق عليها سمة القلق (A-Trait) حالة القلق (A-state) وذلك لقياس القلق لدى تلاميذ المرحلة الابتدائية (الاطفال من سن ٩-١٢ سنة) ويمكن استخدام الاختبار للأصغر سنا والذين يتسمون بقدرة قرائية متوسطة أو فوق المتوسطة، وكذلك مع الكبار ذوى القدرة القرائية الاقل من المتوسط.

ويتكون مقياس حالة القلق من ٢٠ مفردة يسأل فيها الاطفال عما يشعرون به فى لحظة ما، ويتكون مقياس سمة القلق من ٢٠ عبارة ايضا تعكس اجاباتها ما يشعر به الطفل عامة، ويشبه هذا المقياس فى محتواه المقياس الخاص بالكبار وقد صيغت عبارات المقياس بطريقة بسيطة ومناسبة لصغار السن، أما عن شكل الاختبار فقد كتبت عباراته فى ورقة واحدة يشمل الأول منها مقياس حالة القلق (ط -١) ، كذلك أيضا بالنسبة للإختبار الثانى سمة القلق (ط -٢).

وقد صمم مقياس (حالة القلق) لقياس حالات القلق الوقتية Transitory Anx-iety states وهى عبارة عن الشعور المدرك بالحواس من خوف أو توتر أو أنزعاج والتى تختلف كل منها مع الوقت، أما مقياس (سمة القلق) فيقيس الفروق الفردية الثابتة نسبيا والمميزة للشخصية من حيث اختلاف الاطفال فى درجة القلق ويحصل الاطفال ذو سمة القلق المرتفعة على درجات مرتفعة فى مقياس حالة القلق بمقارنتهم بذوى سمة القلق المنخفضة وتستثار «حالة القلق» بالتعرض للمواقف الخطرة أو التهديدية.

وقد قام معد الاختبار بتقنينه على عينة قوامها ٦١٠ تلميذا من الصفوف من الثالث – السادس وتم حساب ثباته بطريقة اعادة التطبيق واستخدام معامل ألفا. والتجزئة النصفية. أما تحديد الصدق فقد استخدم لتحديده صدق المحتوى والصدق المرتبط بالمحكات. كما تم حساب معايير تائية وأخرى مئينية.

المقياس الموضوعى السلوكى لذوى الإعاقة العقلية

Objective Behavioral Assessment of the Mentally Handicapped (OBA)

هذا المقياس اعده فى أصله الاجنبى لارى هاردى وآخرون وعربه وعدله ايمان فؤاد كاشف ، هشام ابراهيم عبد الله ٢٠٠٦ ويعد هذا المقياس اداة لقياس السلوك التكيفى للافراد ذوى الاعاقة العقلية الشديدة والمعتدلة، وهذا المقياس يتميز بالموضوعية ونسبة عالية من الثبات كما انه سهل التطبيق وعبارات المقياس تعكس السلوكيات المحددة المطلوب قياسها وتقع فى خمس فئات هي

١– مهارات رعاية الذات Basic Self Care Skills

٢– المهارات الاجتماعية والمهارات المتقدمة لرعاية الذات Social skills and Advanced Self Care Skills

٣– المهارات المنزلية Sheltered Domestic Skills

٤– مهارات الاتقان الحركى قبل المهنى Pre - vocational Motor Dexter-ity Skills

٥– مهارات اداء العمل Work Performance Skills

وتنقسم المهارات السابقة الى مجموعة اختبارات تضم مهارات فرعية مرتبطة بالانشطة اليومية المألوفة مثل اعداد الوجبات الصغيرة، ترتيب السرير ونظافة حجرم النوم واعداد المائدة للطعام.. الخ

ويتضمن دليل الاختبار وصفا للمهام وطريقة تقدير الدرجات على مقياس تقدير خماسى، كما يشير الى تحديد ثبات المقياس.

مقياس الكفاءة الاجتماعية لاطفال ما قبل المدرسة الابتدائية

اعد المقياس كل من اسماء السرسي، وأمانى عبد المقصود ومراجعة جابر عبد الحميد جابر

يتزايد الاهتمام بالتعرف على نمو طفل ما قبل المدرسة فى الجوانب المختلفة ومنها نموه الاجتماعى والكفاءة الاجتماعية، ويرجع الاهتمام الى أهمية نمو الطفل وشخصيته نموا متكاملا

ومقياس الكفاءة الاجتماعية أعده كل من صموئيل ليفين، وفريمان إلزي، ومـــارى لويز (Lewis, M, Elzey, F, Levine, S فى عـــام ١٩٦٩) تحت اسم كاليفورنيا للكفاءة الاجتماعية لاطفال ما قبل المدرسة (California) (CPSCS) Preschool Social Compeleny Scale)

ويهدف الى قياس كفاءة السلوك البينشخصى لاطفال ما قبل المدرسة وتحديد درجة شعورهم بالمسئولية الاجتماعية وبصورة غير مباشرة قياس مفهوم الاستقلالية.

ويتكون المقياس من ٣٠ مفردة تمثل عينة من انماط السلوك الحرجة فى الوظيفة الاجتماعية لطفل ما قبل المدرسة من سن عامين وستة شهور وحتى الخامسة وستة شهور، وتتم المعايير على اساس تقديرات المعلم للطفل خلال المدى الزمنى عند التحاقة ببرامج الحضانة ورياض الأطفال، وتعطى الفقرات مدى واسعا من الانماط السلوكية مثل الاستجابة للانظمة الروتينية أو الاستجابة للأشياء غير المألوفة واتباع التعليمات، وتقديم تفسيرات، والمشاركة، ومساعدة الاخرين والمبادأة بالانشطة والاستجابة للاحباط وقبول القيود الاجتماعية.

وتتكون كل فقرة من أربع عبارات وصفية تمثل درجات متباينة من الكفاءة بالنسبة للسلوك المقاس وتتطلب الفقرات ملاحظة للأداء الفعلى فى المواقف الطبيعة،

ولكى يستخدم المقياس بكفاءة يجب ان يكون القائم على التقدير على دراية بمضمون المجالات التى يتناولها المقياس.

وقد تضمن دليل المقياس تفصيلات تتطلب من القائم بالملاحظة اتقانها لكل مفردة من مفردات المقياس وكذلك كيفية تصحيح كل مفردة طبقا للمستويات الاربعة (١-٤) والمترتبة على بعضها البعض بمعنى ان الطفل الذى يصل الى المستوى (٤) من الكفاءة يمكن اداء جميع المستويات الاقل فى الكفاءة، وعلى ذلك يتم اختيار مستوى واحد فقط لكل فقرة، وعلى ذلك تكون أقل درجة يحصل عليها الطفل (٣٠) وأعلى درجة (١٢٠) درجة.

وقد تضمن دليل المقياس ايضا خطوات بناء المقياس وتحديد خصائصه السيكومترية فى الولايات المتحدة، كما تضمن أيضا وصفا لخصائص المقياس السيكومترية واجراءات الوصول لها من حيث ترجمته والتحقق من صحتها، كذلك تحديد صدق المقياس فقد تم الاعتماد على صدق المحكمين (١٠) وصدق التكوين (الإتساق الداخلى حيث تم حساب معاملات الإرتباط بين كل فقرة والاخرى وايضا بين كل فقرة والدرجة الكاية والعمر الزمنى والمستوى الاقتصادى لعينات الاطفال من سن ٢,٦ -٥,٦ .

كذلك قامت معدتا المقياس بتحديد ثبات المقياس بطريقة إعادة التطبيق على ٨٠ طفلا بفاصل زمنى اسبوعين وبلغ معامل الثبات ٠,٦٧٥ وهو معامل دال عند مستوى ٠,٠١ كما حسب معامل الثبات بطريقة التجزئة النصفية وبلغ معامل الثبات بعد التصحيح من اثر التجزئة ٠,٨٦٤ وهو معامل دال ايضا عند مستوى ٠,٠١

كذلك تم الاشارة فى دليل الاختبار ان الدرجات الخام التى يحصل عليها المفحوصون يمكن تحويلها الى درجة معيارية (تائية) وتضمن الدليل الدرجات المعيارية المعدلة فى جداول للفئات العمرية المختلفة من سن ٢,٦ الى ٥,٦ للذكور والاناث كل على حدة.

الفصل الخامس عشر

قياس الشخصية عن طريق الأداء والأساليب الإسقاطية

على الرغم من أن اختبارات الأداء والأساليب الإسقاطية تستعمل فى قياس الشخصية منذ أكثر من أربعين عاما فإنها لم تصل إلى درجة النضج التى وصلت إليها الاختبارات والأساليب والطرق التى تكلمنا عنها حتى الآن. لقد كان الحلم الأكبر للمشتغلين بقياس الشخصية أن يصلوا إلى وسائل قياس، تمثل السلوك مباشرة وليس انطباعات متميزة أو تقريرات ذاتية غير دقيقة، ووسائل قياس تعطى نتائج كمية، وتسمح بالمقارنة المباشرة بين الأفراد فى المواقف ذاتها، وكانت اختبارات الأداء والوسائل الإسقاطية - وليس بينهما تمييز قاطع كما يرى كرونباك (٢٧ : ٥٣٩) - كانت ضمن هذه الوسائل، وإن كانت لم تحقق، كما سبق، نضجا يعادل ما وصلت إليه أساليب قياس الشخصية الأخرى.

واختبارات الأداء التى تهدف إلى قياس الشخصية تصمم بحيث تؤدى إلى ظهور دلائل ومؤشرات وعينات من السلوك المميز حقا للشخص، ويمكن أن نحدد الملامح العامة لمثل هذه الاختبارات فيما يلى:

١ -أن يكون الموقف المثير موحدا بقدر الإمكان بالنسبة إلى المفحوصين كلهم.

٢ - يصمم الموقف بحيث يسمح بظهور تنوع فى أنماط السلوك التى يهدف المختبر إلى ملاحظتها.

٣ - يوجه المفحوص إلى الاعتقاد أن خاصية معينة هى التى تقاس بينما يقوم الملاحظ بملاحظة خاصة أخرى أو جانب آخر من السلوك أو الأداء.

٤ - يسجل الملاحظ بدقة طريقة المفحوص فى الأداء وليس كمية السلوك المؤدى.

والمزية الكبرى للملاحظة فى اختبارات الأداء أنها تكشف عن خصائص لا تظهر إلا قليلا جداً فى أنشطة الشخص العادية من هذه الخصائص مثلا: الشجاعة، ورد الفعل للحبوط، وعدم الأمانة، كما أن من مزايا هذا النوع من الملاحظة فى مواقف الأداء أن رغبة المفحوص فى أن يعطى انطباعا حسنا لا تؤثر فى مدى صدق

الاختبار، لأن الشخص بسبب هذه الرغبة سوف يكشف عن خصائصه بدرجة أكبر مما لو لم تكن عنده هذه الرغبة، وعلينا أن نأخذ فى الاعتبار هذا الدافع عندما نفسر نتائج ملاحظاتنا، ومن مزايا اختبارات الأداء أيضًا أنها تسمح إلى درجة بعيدة بمقارنة الأشخاص فى تعرضهم لظروف واقعية موحدة ومتشابهة.

وتتراوح المواقف التى تستخدم الأداء بين المواقف العاملة والمحددة فى بنيتها highly structured والمواقف التى لاتتميز ببنية واضحة محددة Totally un-structured، ويقال أن موقفا ما ذا بنية كاملة ومحددة إذا كان له معنى واحد معين لجميع المفحوصين، بينما يكون الموقف الذى لا تتوفر فيه بنية واضحة محددة متضمنة لعدد قليل جداً من الإشارات، أو ليس فيه سوى نمط أو نظام ضئيل جداً إلى درجة تجعل من الممكن للمفحوص أن يعطيه أى معنى يشاء، ومن أبسط أمثلة المواقف غير محددة البنية الصوت الغريب الذى ينبعث من مكان فى المنزل مغلقا بستائر الظلام سابحا على أمواج السكينة فى وسط الليل.. تسمعه فلا تعرف بالتحديد ما هو؟ ما مصدره؟ يعطيه كل من يسمعه تفسيرا خاصًا، هذا التفسير الخاص يتأثر إلى درجة بعيدة بمخاوفنا، شعورية كانت أم لا شعورية، وباهتماماتنا ومعرفتنا، أى أنه فى الموقف محدد البنية يعرف الشخص بالضبط ما يتوقع منه، وكيف يقوم به، أما فى الموقف غير محدد البنية، فإن الشخص يوجه نفسه ويقودها فى الاستجابة له. وكلما زاد غموض الموقف، أى زادت بنية عدم تحديده زادت الفرصة لتعطى الشخص تفسيرا أو أداء خاصا متميزا له.

وتعتبر المواقف محددة البنية أكثر مناسبة لقياس القدرة، أو قياس أقصى الأداء، لأنها تجبر كل مفحوص على أن يحاول أمرا بعينه، وبينما نجد الأساليب غير محددة البنية، أو الأساليب الإسقاطية، تستعمل ، بعكس مقاييس القدرة، مثيرات لا تتميز غالبا بأية بنية محددة.

وفيما يلى نتناول نوعين من اختبارات الأداء لقياس الشخصية، النوع الأول هو اختبارات المواقف ذات البنية المحددة، والنوع الآخر هو اختبارات المواقف ذات البنية الأقل تحديدا أو الاختبارات الإسقاطية.

أولا: اختبارات الأداء ذات المواقف محددة البنية:

تدخل ضمن هذا النوع من الاختبارات، اختبارات الخلق والإصرار فى العمل،

كذلك الاختبارات التى تقيس الأساليب الإدراكية والمعرفية . perceptual and cog-nitive styles.

لقد كان استعلام تربية الخلق The Character Education Inquiry الذى وضعه هارتشون وماى Hartshorne and May من أكثر المجهودات التى بذلت لتقويم الشخصية عن طريقة أساليب كمية وموضوعية خالصة، وكان اعتقادهما أن سمات الخلق يمكن أن تقاس بدقة عن طريق تعريض الشخص للإغراء فى مواقف يظن أنه يمكن فيها أن يخالف المعايير والقواعد الاخلاقية دون أن يكتشف أمره، وكانت السمات التى درسها هذان الباحثان تتضمن، الصدق، والأمانة فيما يتصل بالنقود، والإصرار والدأب فى العمل، والتعاون، والكرم، (راجع ٣٢، ٣٣، ٣٤).

أما اختبارات الأفراد فى الأساليب والطرق التى يتبعونها فى إدراك وتناول المشكلات، فالواقع أن الدارس الذى يقوم بملاحظة أى سلوك أثناء حل مشكلة يمكن أن يميز بسرعة الاختلافات فى الطرق التى يتبعها المفحوصون فى مواجهة المشكلات وتخطى العقبات. وتعتبر الاختبارات الفردية، مثل اختباروكسلر الذى سبق تناوله، فرصة لملاحظة هذه الأساليب أو العادات المختلفة بين الأفراد، والمعلومات التى نجمعها عن هذه الأساليب الخاصة تساعد فى التنبؤ بنوع المشكلات التى يستطيع الشخص أن يتعامل معها بنجاح، وأى الأخطاء يمكن أن يقع فيها أكثر، وقد تفتح أمامنا طريقا لعلاج مناسب يؤدى إلى تحسين الأداء فى مواجهة المشكلات.

وتهتم اختبارات العمليات المعرفية أساسا بالطريقة التى ينظم بها الشخص المعلومات التى ترد إليه، هذا التنظيم Organising أو إعطاء نمط Patterning كان يمثل الاهتمام الرئيسى عند أصحاب نظرية الجشطالت، ومعظم الاختبارات التى وضعت فى هذا المجال وضعها باحثون ينتمون إلى مدرسة الجشطالت، وقد استعمل هؤلاء الباحثون مثيرات بسيطة لاختبار الجهاز الإدراكى كله معتقدين أن المخ جهاز لتحويل المعلومات وأنه عن طريق مثيرات أو اختبارات إدراكية بسيطة يمكن أن نتعرف على خصائصه المميزة، من هذه المثيرات الإدراكية البسيطة اندماج الخفق flicker fusion والحركة الظاهرة apparant movement فعندما ينظر الشخص إلى ضوء من خلال جهاز دوار يحجب الضوء ثم يكشفه، فإنه يدرك الخفق عندما يكون

الدوران ذا معدلات بطيئة، وكلما زاد معدل الدوران تدريجيا فإن الشخص يستطيع أن يقرر النقطة التى يختفى عندها إدراكه للخفق، وقد وجدت فروق فردية واسعة فى هذه الناحية وهناك من الدلائل ما يشير إلى أن نقطة الاندماج هذه تدل على قدرة الجهاز العصبى على تسجيل تفاصيل المثيرات الواردة، وقد استعملت مقاييس اندماج الخفق بنجاح فى تشخيص إصابات المخ.

أما عن الحركة الظاهرة فالمقصود بها أنه عندما يعرض ضوءان جنبا إلى جنب مع إطفاء أحدهما وإنارة الآخر، بتتابع سريع فإنه يظهر أن الضوء يتحرك قافزا جيئة وذهابا، وفى الحركة الظاهرة هذه، وتسمى أيضًا ظاهرة فاى، نجد أن المخ يقوم بتكامل المثيرات فى نمط، وقد درس كلين وشيليزنجر (٣٨) أثر تغيير الفترات بين الإضائتين، ووجدا أنه كلما قصرت هذه الفترة كانت هناك عتبة فارقة لبدء ظهور الحركة الظاهرة وعتبة فارقة ثانية تختفى عندها، أي يرى الشخص بعدها ضوأين ثابتين، وقد كان الفاصل بين هاتين العتبتين، أى مدى الفترات التى تسمح بظهور انطباع الحركة، أكثر اتساعا عند بعض الأفراد عنه عند غيرهم.

ومن اختبارات الأساليب المعرفية ما يهدف إلى قياس المرونة العقلية والجمود أو التصلب العقلى، فمما يلاحظ كثيرا أن الذين يفشلون فى حل المشكلات يتميزون بالتشتت والوقوف عند الأفكار الخاطئة، فهم مثلا يكررون الدخول فى ذات الممر المسدود فى المتاهة، والواقع أن المواجهة السليمة للمشكلة تتطلب إعادة تنظيم المجالات الإدراكية على أساس المعلومات أو المتطلبات الجديدة، وقد أنشئت اختبارات لقياس درجة المرونة الإدراكية والعقلية، نشير منها إلى اختبار التأهب Einstellung Test (أو اختبار التوجيه orientation أو اختبار الاتجاه العقلى mental set) الذى وضعه لوشينز (٤٤) وهو يعرف باسم اختبار اناء الماء، ويستعمل مشكلات توزيع كميات من المياه فى آنية محددة الاتساع ويجرى الاختبار هكذا مثلا، إذا كان لديك اناء سعته ٧ أرباع جالون وآخر سعته ٤ أرباع جالون فكيف تحصل بالتحديد على ١٠ أرباع جالون؟». وتمشيا مع منطق الاختبار يسمى الاناء ٧ (أ) والاناء ٤ (ب) ويكون الحل هو املأ ب تاركا ثلاثة أرباع جالونات فى أ، أفرغ ب، املأ ب من أ، ثم أملأ أ. وعلى هذا نحصل على ثلاثة أرباع جالونات حسب القاعدة أ – ب، ونحصل على الـ ٧ أرباع جالونات الأخرى من أ، أى أن (أ – ب) + أ = ١٠ وتكون سلسلة المشكلات

كما يلى:

	الآتية	المطلوب			
	أ	ب	جـ		
مثال	٧	٤	–	١٠	٢ أ – ب
١	– ٢١	١٢٧	٣	١٠٠	ب– أ–٢ جـ
٢	– ١٤	١٦٣	٢٥	٩٩	ب– أ–٢ جـ
٣	– ١٨	٤٣	١٠	٢١	ب– أ–٢ جـ
٤	– ٩	٤٢	٦	٢١	ب– أ–٢ جـ
٥	– ٢٠	٥٩	٤	٣١	ب– أ–٢ جـ
٦	– ٢٣	٤٩	٣	٢٠	ب– أ–٢ جـ أو أ– جـ اختبار
٧	– ١٥	٣٩	٣	١٨	ب– أ–٢ جـ أو أ– جـ اختبار
٨	– ٢٨	٧٦	٣	٢٥	أ– جـ أ انطفاء
٩	– ١٨	٤٨	٤	٢٢	ب– أ – ٢ جـ أو أ + جـ اختبار
١٠	– ١٤	٣٦	٨	٦	ب– أ – ٢ جـ أو أ – جـ اختبار

ومن الممكن مساعدة المفحوص فى حل المشكلات الأولى، وتؤدى المشكلات من ١ حتى ٥ التى تحل عن طريق تطبيق قاعدة معينة إلى تكوين (اتجاه عقلى أو تأهب لاستعمال هذه القاعدة أو المعادلة، بعد ذلك تقدم مشكلات الاختبار critical والانطفاء extinction. وفى مشكلة الاختبار كما فى المشكلة رقم ٦ مثلا يمكن أن تؤدى القاعدة أو المعادلة إلى الحل، ولكن هناك طريقة أسهل للحصول على الإجابة، أما فى مشكلة الانطفاء كما فى رقم ٨ مثلا، فإن التأهب، أى استعمال القاعدة لايؤدى إلى الحل ولكن من الممكن بالنسبة إلى المفحوص الذى يتمتع بمرونة أن يكتشف قاعدة أكثر بساطة ولكى يحصل المفحوص على درجة مرتفعة على اختبار التأهب هذا فإن عليه أن يتنبه للمشكلة المباشرة أمامه، مستبعدا أى ذكريات عن الحلول السابقة، ومن المحتمل يكون السلوك غير المرن، المتصلب، الجامد دليلا على عدم القدرة على الفصل بين المعلومات المتعارضة.

ومن الأمثلة الأخرى لاختبارات المرونة الإدراكية اختبارات الأشكال المتضمنة

Embedded Figures test حيث يقدم فيها شكل هندسى غريب ثم يطلب من المفحوص استخراجه من مجال أكثر تعقيدا، وتكون الدرجة هى الوقت الذى يستغرقه المفحوص فى حل مثل هذه المشكلات، وقد أعده للاستخدام فى اللغة العربية سليمان الخضرى الشيخ وأنور محمد الشرقاوى.

كما أن اختبارى الإغلاق الموسيقى السابق عرضهما ينتميان الى نفس النوع من المشكلات (آمال صادق :٢٠٠٧).

ثانيا: اختبارات المواقف غير محدد البنية

ينقسم هذا النوع من الاختبارات إلى ثلاثة أقسام:

الأول: اختبارات حل المشكلات،والثانى، الاختبارات الإدراكية غير محددة البنية، والأخير الاختبارات اللفظية غير محددة البنية، ولنتكلم عن كل قسم من هذه الأقسام.

(أ) **اختبارات حل المشكلات:** من الممكن أن نلاحظ فى اختبار مثل اختبار تصميم المكعبات وبنائها الطريقة التى يتناول بها المفحوص المشكلة أو ما يسمى بطريقة المواجهة method of attack كما يمكن أن نلاحظ رد الفعل للحبوط re-sponse to frustration عند هذا المفحوص كلما واجهته صعوبة أو فشل فى الوصول إلى حل لما أمامه من مشكلات، إلا أننا نستطيع أن نحصل على معلومات أفضل إذا عدلنا الاختبار، وإذا حددنا بالضبط ما نريد ملاحظته.

وقد حدد جولدنر (٣٠) Goldner طريقة المواجهة فى ضوء متغيرين أكثر تحديدا هما: اتجاه الكل ـ الجزء whole-approach والتصلب rigidity وقد استعمل ستة اختبارات منها: اختبار مكعبات معدل، واختبارات تكوين كلمات من عدد معين من الحروف، واختبار الرورشاخ، واختبار الوظيفة، والذى يسأل فيه المفحوص أن يعطى استعمالات أووظائف مختلفة لأشياء معينة مثل الصندوق أو المكنسة أو الورق، وقد أنشأ جولدنر قواعد تصحيح لكل اختبار، فمثلا بالنسبة إلى اختبار الوظيفة تعطى درجة الكل ـ الجزء حسب ما إذا كانت الاستجابة تفتت الشىء إلى أجزاء كأن تكون الإجابة هى مثلا، استعمال الصندوق فى إشعال النار.

ويقيس جولدنر التصلب أيضاً، فمثلا فى اختبار المكعبات يعين التصلب (أ) إذا وجد المفحوص صعوبة فى الحكم على العدد الصحيح من المكعبات و(ب) إذا احتفظ بذات طريقة المواجهة بعد فشلها، أو (جـ) إذا يئس وتوقف دون أن يكمل حل مشكلة كما يميز التصلب بميل المفحوص إلى إعطاء عدد كبير من الاستعمالات والوظائف المتشابهة منطقيا، بينما تتميز المرونة بالتنوع فى الاستعمالات والوظائف.

وإذا قارنا أسلوب جولدنر هذا باختبار اناء الماء الذى سبقت الإشارة إليه نلاحظ أنه فى اختبار اناء الماء نحصل على درجة تمثل خاصية يمكن حصرها، الجمهور مثلا. بينما فى بطارية جولدنر تمثل كل درجة تقديرا لعملية عقلية مفترضة ويمكن أن تظهر فى الأداء بحيث يمكن استنتاج العمليات العقلية منه، وقد وجد جولدنر تأييدا كثيرا من دراساته لفرضه الذى يذهب إلى عمومية السمتين اللتين لاحظهما أولاهما: اتجاه الكل الجزء، التصلب.

ومن اختبارات المشكلات التى تستعمل فى قياس الشخصية اختبارات تكوين المدرك Concept-Formation فقد استعملت فى قياس الشذوذ فى العمليات العقلية. ويمكن أن نأخذ اختبار هنافمان ـ كازانين Hanfmann Kasanin نموذجا لهذه الاختبارات، يستعمل فى هذا الاختبار عدد ٢٢ مكعبا خشبيا ذات ألوان وأشكال وارتفاعات وأحجام مختلفة، توضع على طاولة. ومطبوع على الجانب المختص من كل مكعب مقطع عديم المعنى، وكل مقطع يميز نمطا من المكعبات، مثلا كل المكعبات التى تحمل مقطع م ن ر تكون صغيرة وطويلة، ويخبر الممتحن المفحوص بأن المكعبات أنواع أربعة، نوع منها يسمى م ن ر، وأن عليه أن يكتشف أساس التقسيم ويصنف المكعبات، ويعمل المفحوص على اكتشاف التصنيف كما يجب ، إلا أنه لايسمح له بقلب المكعبات ليرى الأسماء، وبعد كل محاولة للتصنيف يوضح الممتحن خطأ واحدا ثم يطلب من المفحوص أن يقوم بمحاولة أخرى، ويستمر هذا إلى أن يتم حل المشكلة، ويعطى المبدأ الذى يقوم عليه التصنيف. وتوضيح ملاحظة المفحوص ما إذا كان يستعمل فروضا منطقية، مثلا، ربما كانت كل م ن ر مستطيلات، أو يستعمل التخمين، أو يستفيد من التصحيح أو القدرة على استيعاب تأهب غير سليم والتخلص منه وتكوين مدرك جديد. كما يمكن ملاحظة أى أساليب أو طرق شاذة يتبعها المفحوص.

(ب) الاختبارات الإدراكية غير محددة البنية: ويمكن أن نطلق على هذا النوع من الاختبارات الإدراكية، بصفة عامة، الاختبارات الإسقاطية أو الأساليب الإسقاطية Projective Technique في قياس الشخصية، والمميز الرئيسي لهذه الأساليب هي أنها تكلف المفحوص بعمل غير محدد البنية نسبياً، أي عمل يسمح بعدد غير محدد تقريبا من الاستجابات الممكنة، لهذا فإن التعليمات في مثل هذا النوع من الاختبارات تكون عامة بحيث تترك لخيال المفحوص العنان، وكذلك فإن مثيرات الاختبار نفسه عادة ما تكون غامضة وغير محددة، والافتراض الكامن وراء هذه الأساليب أن الطريقة التي يدرك عليها المفحوص مواد الاختبار ويفسرها، أي طريقة بنائه للموقف سوف تعكس الجوانب الأساسية لتكوينه النفسي، أي أن مواد الاختبار سوف تعمل في هذه الحالة كأنها شاشة عرض يسقط عليها الشخص أراءه واتجاهاته وأطماحه ومخاوفه وصراعاته وعدوانيته وهكذا.

وتتميز الأدوات الإسقاطية أيضاً بأنها تمثل طرق اختبار مخفاة discursive لأنه من النادر أن يعلم المفحوص أي نوع من التفسير النفسي سوف يعطي لاستجاباته، كما تتميز هذه الأدوات الإسقاطية بأنها تتخذ اتجاها شاملا global في تقدير الشخصية، ذلك لأنها تركز على الصورة الكاملة للشخصية كلها وليس على قياس سمات منفصلة.

وفيما يلي نتعرض باختصار لأبرز الاختبارات الإدراكية غير محددة البنية.

١ – اختبار بندر ـ جشطالت Bender-Gastalt:

وقد وضعته لوريتا بندر في عام ١٩٣٨ ونشرته تحت اسم (اختبار الجشطالت البصري الحركي واستعمالاته الاكلينيكية)، ويشترك في هذا الاختبار مع بعض الاختبارات التي سبقت الإشارة إليها في أنه نبع من أبحاث مدرسة الجشتالت في الإدراك ونلاحظ أن الاختبارات الإدراكية بصفة عامة تقوم على اختبار كيف يتعرض الشخص للمعلومات ويعيد تنظيمها ثم يذكرها ويقررها، وفي اختبار بندر تعرض أمام المفحوص أشكال ذات أنماط مختلفة من التنظيم، نقط ودوائر صغيرة وخطوط متعرجة متقاطعة وغيرها، ويسأل المختبر المفحوص أن ينقل مجموعة الأشكال، ثم يلاحظ طريقته في الأداء ومدى نجاحه في نقل هذه الأشكال.

وإذا كانت اختبارات الشخصية ذات البناء المحدد مصممة بحيث تقيس سمة

معينة فإن اختبار بندر لايمكن أن نميزه على أساس أنه مقياس لسمة واحدة أو لمجموعة من السمات، ذلك لأن استجابات الأفراد عليه قد تختلف وتتباين بدرجة كبيرة جداً، وهذه الاستجابات يقوم المختبر بملاحظتها ومقارنتها وتفسيرها.

وقد وضعت قواعد لتصحيح اختبار بندر إلا أنه عند استخدام الاختبار فى أغراض تشخيصية لابد أن يعمل مصححه ومفسره على التكامل النوعى لنتائج الاختبار مع أدوات تشخيصيه أخرى.

ومن الممكن معالجة درجات الاختبار إحصائيا وذلك عن طريق ملاحظة العلامات أو الإشارات التى تميز جماعة محك معينة، فمثلا من الممكن تحديد الاستجابات التى تميز العصابيين من العاديين والتى منها مثلا: تكديس الأشكال فى نصف الصفحة فقط، أو العد بصوت مرتفع عند نقل الأشكال وهكذا، وقد أثبتت دراسات أجريت على هذا الاختبار أن هذه الإشارات وغيرها كانت مميزة فعلا بين العاديين والعصابيين (٩).

٢ - **اختبار بقع الحبر لرورشاخ** Rorchach Inkblots

وهو أكثر الاختبارات الإدراكية غير محددة البنية أو الإسقاطية أهمية واستعمالا وتعرضا للمناقشة والجدل، وقد وضعه هرمان رورشاخ Hermann Rorchach وهو سويسرى الجنسية، وفى عام ١٩٢١، وعلى الرغم من أن بقع الحبر كانت تستعمل من قبل فى دراسة التخيل وبعض الوظائف العقلية الأخرى فقد كان رورشاخ أول من استعملها فى البحث التشخيصى للشخصية ككل.

ويتكون الاختبار من عشرة أشكال من بقع الحبر، خمسة منها تتضمن درجات من اللون الرمادى والأسود فقط واثنان باللون الأسود والأحمر، أما الثلاثة الباقية فهى متعددة الألوان، ولسهولة استخدام الاختبار وضعت الصور على بطاقات أبعادها ٧ × ٩,٥ بوصة.

وعند تطبيق الاختبار تقدم للمفحوص بطاقة ثم يسأل أن يتكلم عما يراه، أى عما يمكن أن تمثله بقعة الحبر ويقوم المختبر، إلى جانب تسجيل استجابات المفحوص لكل بطاقة بملاحظة زمن الاستجابة، والوضع الذى يمسك به المفحوص البطاقة، وملاحظات المفحوص التلقائية وتعبيراته الانفعالية، كما يلاحظ أى سلوك عرضى

يصدر عن المفحوص أثناء جلسة الاختبار، وبعد تقديم البطاقات العشر للمفحوص يسأل المختبر المفحوص بطريقة منظمة عن الأجزاء والجوانب المختلفة للبقعة تلك الأجزاء والجوانب التى أعطى لها المفحوص ارتباطات، وكذلك يعطى المفحوص فرصة ليوضح ويشرح استجاباته السابقة، ويعكس الترتيب الذى تقدم به البطاقات رغبة رورشاخ نفسه فى إدخال نظام نفسى يضمن بقاء استثارة المفحوص على مستوى ممكن.

ومنذ وضع رورشاخ اختباره، أدخلت عليه تغييرات طفيفة عن طريق باحثين آخرين، وحدث تحول فى الاهتمام، فبعد أن كان التركيز على التصنيف إلى أنواع أمراض نفسية مختلفة، أصبح التركيز على وصف العوامل الدينامية فى الشخصية Psychodynamics، وقد أحدث بيك (٢٢، ٢٣) Beck، فى الولايات المتحدة، تغييرات بسيطة فى تصحيح الاختبار، ووضع له معايير أولية، إلا أن كلوبفر Klop-fer أحدث تغييرات أساسية فى نظام تصحيح الاختبار، وفى مصر قام سيد غنيم بدراسات استخدم فيها الرورشاخ على عينات، كما أجرى دراسة سابقة عن مدى صلاحية الرورشاخ لقياس الذكاء (٧) ووضع مع هدى براده معايير مصرية للرورشاخ (٧، ٨) وكذلك أجرى فاروق أبو عوف دراسة مقارنة عن مدى الإفادة من الرورشاخ فى التمييز بين حالات العصاب القهرى والأسوياء (١٠)، ودراسة أخرى عن مدى صلاحية الرورشاخ لقياس الابتكارية (١١)، ودراسة ثالثة استخدم فيها المقياس فى التمييز بين القادة وغيرهم (١٢).

ويعتمد تصحيح الاختبار على قواعد لها طابع موضوعى، وهناك عدد من الدرجات الرئيسية التى يمكن أن توضع ضمن الأقسام الرئيسية الثلاثة التالية التحديد المكانى، والمحددات، والمحتوى، وتدل درجات التحديد المكانى عما إذا كانت الاستجابة تستعمل البقعة كلها (W) ، أو إلى قسم فرعى يشيع إدراكه (D) ، أو تفاصيل غير مألوفة أو يندر الإشارة إليها (Dd)، أما المحددات فهى الشكل، أو اللون، أو التظليل التى أخذها المفحوص فى اعتباره، فمثلا تعطى الاستجابة عن الحركة (M) عندما يصف المفحوص أناسا فى حركة و (CF) عندما تعتمد الاستجابة على الشكل واللون، مع اعتبار اللون أكثر أهمية فى تحديد الاستجابة، ويلاحظ المصحح أيضًا مدى ملاءمة الاستجابة لشكل بقعة الحبر، فإذا كانت الأشكال تدرك بطريقة

مألوفة تعطى (F)، وإذا كانت دقيقة بدرجة غير عادية (+F)، أما إذا كانت بعيدة الصلة ببقعة الحبر أو غير محددة ولا مقبولة فتعطى (-F).

أما درجة المحتوى فتسجل ما إذا كانت الاستجابة تشير إلى أشخاص (H) وأجزاء من أشخاص (Hd) أو ملابس (G) وهكذا.

وتوجد درجة لمدى الشيوع (P) وتعطى على أساس تكرار وشيوع استجابات مختلفة بين الناس، أما درجة الأصالة (Q) فتشير إلى الاستجابات النادرة ولكن تعطى تفسيرا مقبولا لبقع الحبر.

ويقوم تحليل استجابات الرورشاخ بعد ذلك على أساس العدد النسبى للاستجابات الواقعة فى الأقسام المختلفة، وكذلك على النسب والعلاقات فيما بين هذه الأقسام المختلفة، ومن أمثلة التفسيرات الكيفية التى شاع استعمالها بالنسبة إلى درجات الرورشاخ ربط استجابات «الكل» بمستوى التفكير، وربط استجابات «اللون» بالانفعالية والحركة الإنسانية بالتخيل وحياة الوهم، وفى الاستعمال العادى للرورشاخ يعطى تأكيد رئيسى على الوصف الكلى للشخص، والذى ينظم الاكلينيكى فيه الناتج فى صورة متكاملة آخذا فى الاعتبار علاقات الدرجات المختلفة بعضها بالبعض الآخر، كما يستفاد فى الواقع من أية معلومات يحصل عليها المختبر عن المفحوص، سواء كانت من اختبارات أخرى أو من المقابلات أو من سجلات تاريخ الحالة، يستفاد من هذا كله فى إعداد الوصف الكامل للشخص.

وفى الإمكان وضع معايير لدرجات الرورشاخ، إلا أنها يجب أن تقوم على جماعات فعلية، وعلى الرغم من أن الرورشاخ قابل للتطبيق ابتداء من سنوات ما قبل المدرسة إلى سن الرشد، فإن معاييره الأساسية كانت مشتقة أساسا من جماعات من الكبار، وقد بذلت جهود حديثة لتوسع مجالات استعمال الاختبار فى مستويات أعمار متباينة (١٨ : ٥٧١).

وقد تعرضت بعض الافتراضات التى يقوم عليها تصحيح الرورشاخ لكثير من النقد والشك نتيجة عدد متزايد من نتائج الأبحاث، فمثلا أثبتت دراسة قورنت فيها بطاقات الرورشاخ فى صورتها الملونة بالبطاقات ذاتها معروضة بحيث تلغى فيها الألوان، أثبتت هذه الدراسة أن اللون ذاته ليس له ذلك التأثير المنسوب إليه على

خصائص الاستجابات على بطاقات الرورشاخ (١٨) كما أشارت دراسات أخرى إلى أن الاستعداد اللغوى يؤثر فى عدد من درجات الرورشاخ كانت تفسر عادة على أنها مؤثرات على سمات معينة للشخصية، وكذلك وجدت بعض الدراسات ارتباطا عاليا بين درجة تعقد استجابات الرورشاخ وبين درجات المفحوصين على اختبار استعداد لغوى، وبين تعقد هذه الاستجابات والسن والمستوى التعليمى (١٨).

ومن العوامل التى تزيد من تعقيد تفسير درجات الرورشاخ مجموع عدد الاستجابات، والذى يعرف «بإنتاجية الاستجابة، أو (R) فمن الصعوبات التى تمثلها إنتاجية الاستجابة مثلا أنه يبدو أن هذه الإنتاجية - كما هو الحال بالنسبة إلى تعقد الاستجابات كما سبق - متصلة بعوامل مثل السن، والمستوى العقلى، والمستوى التعليمى، بل أن الأكثر من هذا مغزى أن هناك ما يشير إلى أن «إنتاجية الاستجابة» تختلف اختلافا دالا من ممتحن إلى آخر (١٨)، وتوحى هذه النتائج كلها بأن إنتاجية الاستجابة والتى تعتبر فى ذاتها محددا لكثير من درجات الرورشاخ متأثرة بعوامل خارجية وبعيدة عن العوامل التى يزعم أن الرورشاخ يقيسها.

أما بالنسبة إلى صدق الاختبار فإن نتائج ما أجرى عليها من دراسات غير مشجعة بصفة عامة على الرغم من كثرتها، إذ أن معظم هذه الدراسات قد فشلت فى إيجاد أية علاقة دالة بين درجات الرورشاخ وبين مجموعات من الدرجات على اختبارات أخرى أو على تقويمات عامة شاملة للشخصية، وقد وجد أن للرورشاخ صدقا تنبؤيا وصدقا مصاحبا ضعيفا عندما درس من حيث علاقته بمحكات مثل التشخيص الإكلينيكى، وبالاستجابة للعلاج النفسى، وبالنجاح أو الفشل فى مهن مختلفة تلعب فيها خصائص الشخصية دوراً مهمًا، ويوجد صراعات ومخاوف وأوهام أمكن تمييزها بصورة مستقلة، وحتى الدراسات التى أعطت نتائج إيجابية فقد اتضح أن فيها أخطاء منهجية خطيرة (٢٧ : ٥٦٤ - ٥٦٥)، ومما ترتب على النتائج السالبة لصدق اختبار الرورشاخ أن تغير التركيز فى التصحيح من التصحيح الإدراكى التقليدى إلى تحليل محتوى الاستجابات، لقد كان رورشاخ ذاته واتباعه يعتمدون على الأسس الإدراكية للارتباط التى يعطيها الشخص لبقع الحبر، وكانت الافتراضات الكامنة وراء هذا الاتجاه هى أن استجابات المفحوص لبطاقات الرورشاخ تدل على استجاباته الإدراكية العادية، وأن سمات الشخصية تؤثر فى الإدراك، وهذا بالنسبة إلى

التقويم النفسى ـــــــ

تعطها أية إشارة إلى الحياة الواقعية أو الأداء الواقعى لهؤلاء العلماء المفحوصين، وهى تشير إلى أن هذه السجلات توضح فى الحقيقة وجود مشكلات جادة فى حياة هؤلاء المفحوصين ولكنها، أى السجلات لاتعطى ما يوضح أن هؤلاء المفحوصين استطاعوا أن يواجهوا هذه المشكلات بنجاح وفاعلية، وأن اتجاه استعمال الرورشاخ فى دراسة هذه الجوانب السوية والقوية فى الشخصية يمثل أيضاً مجالا جديرا باهتمام الباحثين.

٣ – اختبار بقع الحبر لهولتزمان The Holtzmann Inkblot Technique

صمم هذا الاختبار على نمط الرورشاخ، إلا أن هولتزمان (٢٧) حاول فيه أن يتجنب العيوب الفنية الرئيسية للرورشاخ، والتغيرات التى أدخلها هولتزمان على الاختبار سواء فى مواد المثيرات أو الطريقة تغيرات كبيرة وشاملة تجعل من الممكن أن ندرسه ونقومه دون الإشارة إلى الخصائص النظرية أو العملية للرورشاخ، ويتضمن اختبار هولتزمان مجموعتين متماثلتين من سلاسل البطاقات، تضم كل مجموعة عدد ٤٥ بطاقة، والمطلوب من المفحوص أن يعطى استجابة واحدة لكل بطاقة وعلى هذا تثبت إنتاجية الاستجابة بالنسبة إلى كل مفحوص ونتلافى كثيرا من عيوب التصحيح التقليدى للرورشاخ وطريقة إجراء الاختبار وتصحيحه مقننة تقنينا جيدا ومشروحة بشكل واضح، وللاختبار معايير لعينات سوية نبدأ أعمارها من الخامسة إلى الرشد، وعينات غير سوية، ويبدو ثبات المصحح مقبولا، ونتائج الثبات ، سواء الاتساق الداخلى، أوالتكافؤ مشجعة ، وإن كانت دراسات الصدق، مازالت قليلة، وللاختبار صورة جماعية تعطى درجات قريبة من الدرجات التى يعطيها الاختبار عندما يطبق بصورة فردية.

٤ – اختبار تفهم الموضوع Thematic Apperception Test (TAT)

وهناك اختبار آخر من الاختبارات الإدراكية غير محددة البنية، أو الإسقاطية، إلا أنه يختلف عن الاختبارات التى سبقت مناقشتها فى أن مثيراته تعتبر أكثر تحديدا فى البنية نسبيا، ولأنه يتطلب استجابات وأنشطة أكثر تعقيدا وأكثر خضوعاً للتحكم العقلى من مثيلاته من الاختبارات السابقة، ويعتمد تفسير هذه الاستجابات عادة على تحليل محتوى ذى طابع كيفى وذلك تمشيا مع الاتجاه الغالب إلا أنه من الأكثر فائدة أن تفسر الاختبارات الإسقاطية عامة على أساس تحليل المحتوى، ويعطى اختبار تفهم

الموضوع وغيره من الاختبارات التى تعتمد على القصص مجالات أكثر اتساعا لتحليل المحتوى من اختبار الرورشاخ.

وكان أول ما نشر عن اختبار تفهم الموضوع مقال نشره هنرى موري H. Mur-ray وزميله Morgan وزميلته فى عام ١٩٣٥، واستمرت الأبحاث التى قام بها موري وزملاؤه فى العيادة النفسية فى جامعة هارفارد بالولايات المتحدة ثم ظهرت فى كتاب فى عام ١٩٣٨ (٥٠) وقد استعمل هذا الاختبار منذ ظهوره استعمالا واسعا فى الممارسة الإكلينيكية وفى البحث، كما يعتبر نموذجا اقتدى به كثير من اختبارات الشخصية.

تتكون مواد الاختبار من ١٩ بطاقة تحتوى كل منها صورة غامضة باللونين الأبيض والأسود، بالإضافة إلى بطاقة خالية من أى صورة، ويطلب من المفحوص أن يكون قصة تناسب كل صورة، ذاكرا ما أدى إلى ما تمثله الصورة، واصفا ما يحدث وما تحس به وتفكر فيه الشخصيات الممثلة فى الصورة، وعليه أن يذكر النتيجة التى يتوقع أن ينتهى إليها ما يصفه، أما بالنسبة إلى البطاقة البيضاء التى ليس عليها أى صورة، فإن التعليمات تطالبه بأن يتخيل صورة ما على البطاقة ويصفها ثم يذكر قصة عنها . وتتطلب الطريقة الأصلية كما حددها موري وجاستين لكل منهما ساعة، مع استعمال ١٠ بطاقات فى كل جلسة، ونختار البطاقات المحجوزة للجلسة الثانية على أساس أنها غير مألوفة وشاذة وذات طابع درامى، والتعليمات المصاحبة فى أن يطلق المفحوص لخياله العنان وتوجد أربعة مجموعات من الاختبار، وهى متداخلة، وتتكون كل مجموعة من ٢٠ بطاقة، منها مجموعة للأولاد، وأخرى للبنات، وثالثة للذكور فوق سن ١٤ ، ورابعة للإناث فوق سن ١٤، ويستعمل معظم الاكلينيكيين مجموعة مختصرة من بطاقات منتقاة، ونادرا ما يعطى أكثر من عشر بطاقات لمفحوص واحد.

وعلى الرغم من أن الاختبار أصلا هو اختبار فردى شفوى يعطى فى الموقف الاكلينيكى، فمن الممكن إعطاؤه كتابة، أى يقوم المفحوص فيه بالكتابة بدل الكلام كما يمكن اعطاؤه فى جماعة، وتشير نتائج بعض البحوث إلى أنه فى حالة إعطائه جماعيا وكتابة يكون إنتاج المواد ذات المعنى والمغزى أكثر سهولة ويسرا (١٨ :

٥٠١) وهناك ميل إلى تأكيد أهمية العوامل والمؤثرات الاجتماعية فى الموقف الاختبارى بالنسبة إلى اختبار تفهم الموضوع. كما لاحظنا أثر هذه العوامل والمؤثرات ونوهنا بها بالنسبة إلى اختبار الرورشاخ، ونكرر أن هذا أيضاً مجال طيب لمزيد من البحث النفسى الاجتماعى.

أما فى تفسير قصص اختبار تفهم الموضوع يحدد المختبر أولا البطل hero وهو عبارة عن الشخصية ، ذكرا كانت أم أنثى، التى يتوحد معها المفحوص ثم يحلل محتوى القصص أساسا وفق ما حدده من حاجات needs وضغط Press ، وقد سبقت الإشارة إلى الحاجات عندموری عند الكلام عن مقياس التفضيل الشخصى الذى وضعه ادواردز، ومن أمثلة هذه الحاجات الإنجاز achievement والاعتداء aggres-sion ، والإعالة nurturance ، والجنس، إلا أن موری يعطى حاجات معينة أهمية كبيرة فى التفسير مثل: تجنب الأذى، البحث عن الكمال، الاعتداء، السيطرة، العدوان الموجه نحو الذات ، الاستنجاد، السلبية، الجنس (٩ : ١٤٤) وما بعدها.

ومع معرفة البطل الرئيسى، ودراسة الحاجات والدوافع الأساسية للسلوك لابد من معرفة ما أسماه موری الضغط أوالضغوط، وهى تشير إلى القوى البيئية التى قد تيسر أو تعوق إشباع الحاجات، ومن أمثلة هذه الضغوط ما يشار إليه من التعرض للهجوم أو النقد بواسطة شخص آخر، وتقبل العطف والمحبة، والتعرض للخطر المادى أو الجروح المادية.

وعند تقدير أهمية أو قوة حاجة أو ضغط معين عند الشخص يعطى اهتمام خاص بعمق وحدة وتكرار حدوث هذه الحاجة أو الضغط فى القصص المختلفة بالإضافة إلى الاهتمام بارتباط الحاجة أو الضغط بصورة معينة بصفة خاصة، والافتراض القائم هو أن أية استجابات مخالفة للاستجابات الشائعة لكل صورة يغلب أن يكون لها - أى الاستجابات المخالفة - مغزى بالنسبة إلى المفحوص.

وعلى الرغم من أن تفسير اختبار تفهم الموضوع كان أساسا تفسيراً كيفيا وانطباعيا كما سبقت الإشارة ، فإن كرونباك (٢٧ : ٥٧١) يذهب إلى أنه من الممكن إنشاء وتطوير نظم موضوعية لتفسير هذا الاختبار، وفى رأيه أن هناك عدداً كبيراً من المتغيرات العامة المشتركة التى يمكن أن نلاحظ قوتها تقريبا فى كل أداء على

الاختبار، ومن أمثلتها، إدراك السلطة، ورد الفعل للأعمال الصعبة جداً، والأصالة، والاعتماد على الحظ وتأثير السحر، وتوجد أمثلة متعددة لمحاولات التصحيح الكمية التى أعطت ثباتا عاليا للمصححين، إلا أن تطبيق هذه الأساليب يستغرق وقتا طويلا لهذا ندر استعمالها فى المجال الاكلينيكى.

وقد نشر قدر لابأس به من المعلومات المعيارية عن الاختبار فيما يتصل بأكثر الاستجابات المميزة لكل بطاقة شيوعا، وبالطريقة التى تدرك بها كل بطاقة، والموضوعات المذكورة، والأدوات التى تنسب إلى الشخصيات، وسرعة الاستجابات، وطول القصص، وعلى الرغم من أن هذه البيانات الخاصة بالمعايير تمثل إطارا عاما لتفسير الاستجابات المفردة، فإن معظم الاكلينيكيين يعتمدون اعتمادا رئيسيا على «معايير ذاتية» بنوها من خبرتهم الشخصية فى إعطاء الاختبار.

وقد استعمل اختبار تفهم الموضوع استعمالا واسعا فى بحوث الشخصية، ومن أمثلة استعماله فى مصر بحث أحمد عبد العزيز سلامة (٣) . وركز بعض هذه البحوث على الافتراضات التى تقوم عليها تفسيرات الاختبار مثل توحد الذات مع البطل، والمغزى الذاتى للاستجابات غير المألوفة. وعلى الرغم من أن هذه البحوث لا تعطى تأييدا للصدق التنبؤى والصدق التلازمى، فإنها تضيف إلى صدق التكوين الفرضى لتفسيرات اختبار تفهم الموضوع. ومن الافتراضات الأساسية الأخرى التى يشارك فيها هذا الاختبار الاختبارات الإسقاطية الأخرى أن ظروف المفحوص العاطفية والدافعية وقت الاختبار تؤثر على استجابات المفحوص لموقف الاختبار غير محدد البنية، ويتوافر قدر كبير من نتائج الدراسات التجريبية يشير إلى أن ظروفا مثل الجوع، والحرمان من النوم، والحبوط الاجتماعى، ومعاناة للفشل فى موقف اختبارى سابق، كلها تؤثر على استجابات اختبارتفهم الموضوع، وبينما يؤكد هذا صحة الفرض الإسقاطى فإن حساسية هذا الاختبار لهذه العوامل المؤقتة يزيد من صعوبة الكشف عن سمات الشخصية الثابتة (١٨ : ٥٠١).

ولعلنا نلاحظ أن اختبار تفهم الموضوع قد صمم أساسًا بحيث يشمل الجوانب الشخصية جميعا، كل الأفكار والمشاعر والسلوك، وعلى هذا فلا يستطيع أن يسير موضوعا أو جانبا واحدا بعمق، ولهذا فقد أنشئت اختبارات تفهم موضوع تركز على

جانب بعينه، أى أنها تصمم لاستدعاء استجابات تدور كلها حول مسألة واحدة وقد بدأها موري وليكرت Murray and Likert فى دراسة بين العمال والإدارة وذلك عن طريق تقديم صورة تمثل تظاهرات مع رجال الأمن.

أما المحاولة المستفيضة المتعمقة للاستفادة من هذه الاختبارات المركزة فى تفهم الموضوع فكانت محاولة ماكيلاند ورفاقه (٤٧، ٤٨) وكذلك محاولة أتكنسن (١٩، ٢٠).

لقد اختار ماكيلاند أربع صور، منها اثنتان من اختبار تفهم الموضوع كان القصد استدعاء الاتجاهات نحو الإنجاز. وكانت هذه البطاقات الأربع تمثل موقف عمل (رجل أمام ماكينة)، وموقف مذاكرة (تلميذ يجلس إلى مكتب ومعه كتاب)، وصورة أب وابن، وصورة ولد يبدو مستغرقا فى أحلام اليقظة، ويقاس دافع الإنجاز فى كل قصة تعبر عن التنافس والكفاح للوصول إلى مستوى أو تحقيق هدف أما أتكنسن فقد استعمل اختبار آخر هو French Test of Insights ولذات الغرض، أى قياس الدافع إلى الإنجاز، ويقوم الاختبار على إعطاء وصف موجز للسلوك مثل: «يعطى الفتى «بل» فرصته لزميله ليفوز»، ويسأل المفحوص أن يعطى تفسيرا لهذا، ويتضمن الاختبار عددا من مثل هذه العبارة وتتكون الدرجة عليه من عدد المرات التى تذكر فيها الرغبة فى الإنجاز باعتبارها دافعا، وتمثل هذه الدراسات خطا من الأبحاث الغنية التى نبعت أصلا من اختبار تفهم الموضوع.

٥ - اختبار تفهم الموضوع للأطفال:

Children's Apperception Test (CAT)

وعلى الرغم من أن اختبار تفهم الموضوع الأصلى وضع بحيث يصلح فى التطبيق على أطفال يبلغون من العمر أربع سنوات فإن اختبار تفهم الموضوع للأطفال قد صمم خاصة ليطبق على الأطفال ممن تتراوح أعمارهم بين ثلاث وعشر سنوات، ويستعمل هذا الاختبار على بطاقات صور الحيوانات بدل البشر على أساس افتراض أن الأطفال يسقطون على الحيوانات بسهولة أكثر من إسقاطهم على البشر، إلا أن صور الحيوانات موضوعة فى مواقف بشرية، وهذه الصور مصممة بحيث تستدعى عند الأطفال أوهاما وتخيلات تتصل بمشكلات التغذية والأنشطة الفنية الأخرى،

ومشكلات العلاقات بين الأشقاء، والعلاقات مع الأبوين، والعدوان، وعمليات الإخراج وضبطها، والخبرات الأخرى الخاصة بالطفولة، وعلى العكس من افتراض واضعى هذا الاختبار، أثبتت دراسات متعددة أن استعمال الصور البشرية وصور الحيوانات لم يعط اختلافات تذكر، بل ظهر غالبا أنه فى حالة استعمال الصور البشرية كانت إنتاجية الأطفال من المواد ذات المغزى الاكلينيكى أكبر منها فى حالة استعمال صور الحيوانات، وفى ضوء هذه النتائج صمم واضعوا الاختبار صورة بشرية منه، هى CAT-H للاستعمال مع الأطفال الكبار، وخاصة من يزيد عمرهم العقلى عن عشر سنوات، وذهبوا إلى أن أيا من الصورتين، البشرية والحيوانية – يكون أكثر فاعلية حسب سن الطفل وخصائص شخصيته.

(جـ) الاختبارات اللفظية غير محددة البنية:

على الرغم من الاختبارات غير محددة البنية التى نوقشت حتى الآن تتطلب استجابات لفظية بينما المثيرات غير لفظية، فإن هناك نوعا من الاختبارات غير محددة البنية أو الإسقاطية تعتمد على مثيرات لفظية كما تتطلب استجابات لفظية سواء كانت منطوقة أو مكتوبة، وهى مناسبة للتطبيق الجماعى بالكتابة، ونذكر فيما يلى مثالين من هذه الاختبارات.

١ – اختبارات تكملة الجمل:

وتستعمل اختبارات تكملة الجمل فى اختبارات الاستعداد اللغوى إلا أن اختبارات تكملة الجمل ذات البنية غير المحددة تختلف عن الأولى فى أنها تتطلب استجابات وتكملات متنوعة ومتباينة على أساس أنها نوع من الاختبارات الإسقاطية، وغالبا ما تعطى الكلمة الأولى، أو بدايات الجمل، وعلى المفحوص أن يكتب النهاية ، ومن أمثلتها ما يلى:

أنا أحس..

ما يضايقنى ..

النساء ..

وتصاغ بدايات الجمل بحيث تستدعى استجابات تتصل بمجال الشخصية المراد التعرف عليه، وتعتبر هذه المرونة التى يتمتع بها أسلوب تكملة الجمل بميزة هامة

تزكيه للاستعمال الاكلينيكى وفى البحث، وقد نشرت صور مقننة من هذه الطريقة للاستعمال فى مجال أوسع، منها قائمة روتر للجمل الناقصة Rotter Incomplete Sentense Blank.

وتتكون من عدد ٤٠ بداية للجمل، وتنص التعليمات الموجهة إلى المفحوص على أن يكمل هذه الجمل ليعبر عن مشاعره الحقيقية، وألا يترك أية جملة ويتأكد من أنه يكون جملة كاملة، ومن أمثلة هذه الجمل الناقصة.

أنا أحب ..

أن خوفى الأكبر..

أحيانا...

أنا أرغب..

والدى...

وتقدر كل تكملة على مقياس ذى سبع نقط حسب درجة التوافق أو عدم التوافق التى تدل عليها التكملة، ويعطى دليل الإختبار عينة من التكملات والدرجة المناسبة لها مما يسمح بقدر كاف ومناسب من التقدير الموضوعى للدرجة ويمثل مجموع التقديرات درجة كلية للتوافق يمكن استعمالها فى اغراض الانتقاء، ومن الممكن أن يفحص محتوى الاستجابة للتعرف على ما قد يكون فيها من مؤشرات ذات دلالة تشخيصية، وتشير دراسات صدق الاختبار إلى إمكاناته الواسعة فى الاستعمال.

٢ - **اختبار تداعى الكلمات**: استعمل تداعى الكلمات فى مراحل علم النفس التجريبى والقياس العقلى المبكرة للتعرف على عمليات التفكير، ثم اتسع مجاله فى الميدان الاكلينيكى عندما انتشر أسلوب التحليل النفسى ولعل من أهم الاختبارات المعتمدة على تداعى الكلمات والتى تزايدت أهميتها منذ وضع اختبار كنت - روزانوف للتداعى الحر Kent-Rosanoff Free Association Test والذى وضع فى عام ١٩١٠، واعتمد تماما على التصحيح الموضوعى وعلى المعايير الإحصائية، وتتكون الكلمات المثيرة من مائة كلمة محايدة شائعة اختيرت لأنها تميل إلى استدعاء مثيرات متشابهة عند الناس بصفة عامة. ومثلا يميل معظم الناس الى ان يستجيبوا لكلمة طاولة بكلمة كرسى ولكملة ظلام بكلمة نور، وأعدت مجموعة من جداول

التكرار- واحد لكل كلمة مثير- مبينة عدد المرات التى أعطيت فيها الاستجابة لعينة تقنين عددها ألف راشد سوى.

وفى تصحيح اختبار كنت- روزانوف، تستعمل قيمة وسيط التكرار للاستجابات التى أعطاها المفحوص للمائة كلمة المثيرة على أنها «مؤشر للشيوع»، وتعتبر الاستجابات التى لاتوجد فى جداول المعايير «فردا» وقد بينت مقارنات الذهانيين بالأسوياء أن الذهانيين أعطوا عدد استجابات «فرد أكثر كما حصلوا على «مؤشر للشيوع» منخفض عن الأسوياء. إلا أن القيمة التشخيصية للاختبار بدأت تتناقص مع التحقق التدريجى من أن تكرار الاستجابات يختلف باختلاف السن، والمستوى الاجتماعى الاقتصادى والمستوى التعليمى وغيرها من العوامل، وعلى هذا فإن التفسير السليم للدرجات على الاختبار يتطلب معايير الجماعات الفرعية، وكذلك معلومات إضافية عن المفحوص، وعلى أية حال فإن الاهتمام بهذا الاختبار مستمر ومتجدد فى مجالات السلوك اللغوى والشخصية فضلا عن استعمالاته الاكلينيكية. وقد قام فؤاد أبو حطب نقل هذا الاختبار إلى اللغة العربية وطبقه على عينات مصرية بلغ حجمها ١٠٠٠ مفحوص وأعد له معايير تقبل المقارنة مع المعايير الأجنبية(*).

(*) فؤاد أبو حطب: المعايير المصرية لاختبار تداعى الكلمات: فى بحوث تقنين الاختبارات النفسية، المجلد الأول (تحرير فؤاد أبو حطب)، القاهرة، مكتبة الأنجلو المصرية، ١٩٧٧.

الفصل السادس عشر

قياس الميول

يرجع الاهتمام بالدراسة النفسية للميول إلى التوجيه التربوى والمهنى كما يرتبط إنشاء الاختبارات التى تقيس هذه الميول بعمليتى الانتقاء والتصنيف المهنى، لأن للميول أهمية عملية، ويجب أن يؤخذ فى الاعتبار من جانب كل من طالب العمل وصاحب العمل، فالأداء فى عمل من الأعمال هو خلاصة للاستعداد وللميل معا، وعلى هذا فإن قياس هذين المتغيرين، أى الاستعداد والميل، يسمح بدرجة أكبر من الدقة فى التنبؤ بالأداء من الاقتصار على واحد منهما فقط. ذلك لأنهما، أى الاستعداد والميل، على الرغم من ارتباطهما فإن مستوى مرتفعا فى أحدهما ليس يعنى بالضرورة مستوى مرتفعا فى الآخر، فقد يكون عندالشخص استعداد للنجاح فى نمط معين من النشاط، سواء كان دراسيا، أو مهنيا، أو ترفيهيا، دون أن يكون عنده الميل المناسب لهذا النشاط أو قد يكون عنده ميل لعمل ما ولكنه يفتقد الاستعدادات المطلوبة لهذا العمل (راجع ١ : ٢٧٥ وما بعدها)، أما وقد تكلمنا من قبل عن قياس الاستعدادات، فإننا نتناول فى هذا الفصل قياس هذا المتغير الآخر وهو الميل.

مقاييس الميول

يعتبر اختبار الميول استفتاء أو استبيانا مطولا يستخدم أسلوب التقرير الذاتى الذى تكلمنا عنه فى الفصل الثالث عشر ويهدف هذا الاستبيان إلى الحصول على معلومات بأن نجعل الشخص يصف خصائصه ومميزاته الذاتية وهذا الاختبار أو الاستبيان عبارة عن نوع من المقابلة المكتوبة، ولما كان يتضمن عدداً كبيراً من الأسئلة غير المباشرة، فإنه يفضل المقابلة الشفوية المباشرة.

والواقع أن طريقة السؤال المباشر للمفحوص لتتعرف على ميله تعنى أن نسأله عما يميل إليه من أعمال أو دراسات أو ترويح، وهذه الطريقة تعرضت لنقد شديد فى وقت مبكر، فى خلال العقد الثانى من القرن العشرين. فقد تجمعت دراسات شتى بينت أن الإجابة عن الأسئلة المباشرة فيما يتصل بالميول كثيرا ما تكون غير ثابتة. وسطحية، وغير واقعية (١٨)، وترجع أسباب هذا إلى أن معظم الناس ليس لديهم المعلومات الكافية عن الأعمال ومواد الدراسة والأنشطة المختلفة المتضمنة فى عمل

ما، ومن ثم فإنهم يكونون غير قادرين على أن يقروا ما إذا كانوا سوف يحبون كل ما يشتمل عليه ما يحبونه من نشاط، وقد يكون ميلهم، أو عدم ميلهم، بالنسبة إلى عمل من الأعمال راجعا إلى قصور ما لديهم من معلومات عما يتضمنه هذا العمل من تفاصيل وإجراءات وتصرفات تجرى كل يوم، ومن أسباب عدم الثبات والسطحية وعدم الواقعية هذه شيوع صور وأفكار جامدة ثابتة، عن بعض الأعمال والمهن، كما هو الحال بالنسبة إلى المدرس والمهندس والمحامى والطبيب، وغيرهم، هذه الصور والأفكار تشيع وخاصة بين الشباب، عن طريق وسائل الإعلام المختلفة، فالمشكلة إذن هى أن الفرد ليس فى وضع يسمح له بأن يعرف ميوله فى المجالات المختلفة قبل أن ينخرط فى سلك العاملين فى هذه المجالات، وقبل أن تفوت الفرصة للاستفادة من معرفته حين لا يستطيع تغيير عمله أو تبديله، لهذا كله عدل عن طريقة السؤال المباشر وتحول الاتجاه إلى الاعتماد على الأسئلة غير المباشرة المتضمنة فى اختبارات الميول التى سبق وصفها بأنها بمثابة المقابلة المكتوبة.

وقد اتبعت فى بناء اختبارات الميول وتصحيحها عدة أساليب.

١ – فمنها ما اعتمد فى بناء تكوينه على أساس تجريبى emperical واقعى، مثل اختبار سترونج للميول المهنية Strong Vocational Interest Blank (SVIB).

٢ – ومنها ما اعتمد على أساس التنوع والتعدد، مثل اختبار كيودر Kuder Preference Record

٣ – ومنها أخيرا ما اعتمد على أساس منطقى مثل اختبار لى ـ ثورب The Lee-Thorpe Inventory

وسوف نتناول كلا من هذه الاختبارات.

أولا: اختبار سترونج للميول المهنية:

نشأ اختبار سترونج مع مجموعة أخرى من اختبارات الميول التى انبثقت عن سيمنار انعقد فى العام الدراسى ١٩١٩ – ١٩٢٠ فى مؤسسة كارينجى للتكنولوجيا فى الولايات المتحدة Carnegie Institute of Technology. حيث كان موضوع هذا السيمنار هو الميول، إلا أنه لم يكتب لأى من تلك الاختبارات الاستمرارية والازدهار

مثل ما كتب لستورنج الذى تعرض منذ نشأته للمراجعة والاتساع والبحث، وقد قام بترجمته إلى اللغة العربية عطية محمود هنا.

وتتكون الصورة الحالية لاختبار سترونج، فى أصله الأجنبى، والمنشور فى عام ١٩٦٦ من عدد ٢٩٩ عنصرا مجمعة فى ثمانية أجزاء، وفى الأجزاء الخمسة الأولى يسجل المفحوص تفضيله بأن يضع دائرة حول أحد الحروف L (يحب Like) أوا (محايد Indiffirent) أو D (لا يحب Dislike) بالترتيب، وكل واحد من هذه الأجزاء الخمسة مختص بواحد من الأقسام التالية: المهن، المواد الدراسية، الأنشطة المسلية، وأنشطة متنوعة (مثل اصلاح الساعات، أو التعبير عن الآراء والأحكام علنا بصرف النظر من نقد الآخرين...) والخواص الغريبة وغير المألوفة للناس، أما الأجزاء الثلاثة الأخرى من الاختبار فتتطلب من المفحوص أن يرتب أنشطة معينة حسب تفضيله، ويقارن ميله فى أزواج من العناصر، وأن يقدر قدراته الحالية وخصائصه ومميزاته.

ويصحح الاختبار بمفتاح خاص لكل مهنة، وحاليا يوجد ٥٤ مقياسا مهنيا لتصحيح الصورة الخاصة بالرجال من الاختبار و٣٢ مقياسا مهنيا لتصحيح الصورة الخاصة بالنساء، وتنشأ مقاييس جديدة من وقت لآخر كلما تجمعت معلومات وبيانات عن جماعات عمل أو مهن أخرى. وعند إنشاء هذه المقاييس المهنية قورنت استجابات أشخاص ناجحين فى مهنة باستجابات الرجال بصفة عامة أو النساء بصفة عامة، وتتكون جماعة المرجع هذه على أساس افتراض هو أن معظم مقاييس المهن فى اختبار سترونج تتصل بالمهن والوظائف التجارية العليا، إلا أنه قد ثبت أن استخدام جماعة مرجع ممثلة بكل مجتمع الذكور كان أقل نجاحا فى التمييز بين المهن المفردة، فميول المهنيين ورجال الأعمال كجماعة المرجع العامة أدى هذا إلى طمس وعدم وضوح الفروق بين المهن ذات المستوى المرتفع.

وقد اختبرت عناصر الاختبار لكل مقياس مهنى وأعطيت لها أوزان على أساس اختلاف تكرار اختيار الرجال فى المهنة المتصلة بالمقياس والرجال بصفة عامة، فمثلا فى مقياس المحامى يدل وزن + ١ على أن الاستجابة تعطى أكثر فيما بين المحامين عنها بين الرجال عموما، بيلما يدل وزن -١ على أن الاستجابة تعطى أقل قيما بين المحامين عنها بين الرجال عموما، والاستجابات التى لاتميز بين المحامين والرجال

بصفة عامة، لاتظهر فى مقياس المحامى. بصرف النظر عن مقدار اختيار المحامين لها، وتتكون الدرجة الخام للمفحوص على كل مقياس مهنى من المجموع الجبرى لأوزانه الموجبة والسالبة وتحول هذه الدرجات الخام إلى درجات معيارية متوسطها ٥٠ وانحرافها المعيارى ١٠ حسب توزيع الدرجات فى كل جماعة مهنية -ocupa tional criterion group وجماعات المحك المهنية هذه كانت تتكون عادة من ٣٠٠ شخصًا من الرجال تتراوح أعمارهم بين ٢٥ و ٥٥، ملتحقين بأعمالهم لمدة ثلاث سنوات أو أكثر، وقد قرروا أنهم راضون عن أعمالهم.

وقد أعطيت تقديرات حروف لزيادة سهولة وتفسير الدرجات المعيارية، فمثلا التقدير (أ) (A) يمثل درجة عند أو أكبر من $\frac{1}{2}$ ع (انحراف معيارى) أي نقطة نحو ٦٩٪ من جماعة المهنة المعينة قد أعطوا استجاباتهم، وبالمثل فإن الحد الأدنى لتقديرات ب (B) قد وضع عند -٢ ع، وتقدر الدرجات أقل من -٢ ع بالتقدير جـ (C). وعلى هذا فإن التقدير جـ يعنى أن الشخص قد حصل على درجة على هذا المقياس بالذات أقل من ٩٨٪ من العينة المماثلة، وعلى أى حال فإنه جدير بالملاحظة أنه مهما كانت الصورة التى تمثل بها الدرجات على اختبار سترونج، فإن هذه الدرجات معتمدة على الاختلافات بين ميول جماعة مهنية وميول الرجال عموما، وعلى هذا فإن درجة مرتفعة على مقياس الطبيب مثلا تدل على أن ميول الشخص تختلف عن ميول الرجال بصفة عامة فى نفس الاتجاه الذى تختلف فيه ميول الأطباء.

هذا هو المقصود من قولنا أن سترونج قد بنى اختباره على أساس تجريبى عملى صرف (٢٧ : ٤٠٦)، دون أن يعتمد إلا على قدرضئيل جداً من المسلمات والافتراضات النفسية، وأقام تصحيحه للاختبار - كما لاحظنا - على أساس الارتباطات الملاحظة للاستجابات مع المحك.

ومن الممكن أن يصحح اختبار سترونج بالنسبة إلى الفرد لمهنة واحدة: فمثلا عندما تهدف إلى الانتقاء فإننا نحب أن نعرف إلى أى مدى تتشابه ميول طالب الوظيفة مع أولئك الناجحين من العاملين فيها، إلا أن الاختبار غالبا ما يصحح كله مما يعطى فرصة لدراسة النمط الكلى لميول المفحوص، ويعطى هذا النوع من التحليل فرصة أوثق للتنبؤ المهنى الغالب.

ويتضمن اختبار سترونج، إلى جانب المقاييس المهنية، أربع مقاييس غير مهنية، هى:

مستوى التخصص	Specialization Level (SL)
ومستوى المهنية	Occupational Level (OL)
والذكورة والأنوثة	Masculanity - Feminity (MF)
والتحصيل الأكاديمى	Academic Achievement (AACH)

وقد أنشئ مقياس مستوى التخصص بمقارنة استجابات المتخصصين فى الطب باستجابات جماعة من الأطباء عامة، وقد وجد أن هذا المقياس يمكن تطبيقه فى مجالات أخرى غير الطب فى التعرف على أولئك الذين سوف يجدون متعة فى الدراسة المتقدمة التى تتضمن تخصصاً دقيقا، أما مقياس مستوى المهنة فيقيس الاختلاف بين ميول العمال غير المهرة فى ناحية ورجال الأعمال والمهنيين فى الناحية الأخرى، أما مقياس الذكورة والأنوثة فإنه يوضح درجة التشابه بين ميول الشخص وميول الرجال والنساء.

أما مقياس التحصيل الأكاديمى، وهو أحدث إضافة إلى المقاييس غير المهنية، فقد أنشئ بمقارنة استجابات جماعات من طلاب المدارس الثانوية والجامعة ممن حصلوا على درجات مرتفعة بجماعات ممن حصلوا على درجات منخفضة، وقد تبين من البحوث التى تلت إضافة هذا المقياس أنه يعكس ميولا للاجتهادات العلمية والعقلية أكثر من ميول للبيع أو الأعمال التجارية (١٨ : ٤٧٠).

وقد تعرض اختبار سترونج لبرنامج مستمر من البحث الذى أعطى قدراً كبيرا من البيانات المتصلة بثباته وصدقه فقد كانت معاملات الثبات - بالتجزئة النصفية على أساس أرقام العناصر الزوجية والفردية - لكل مقياس فرد ٠,٨٨، وكان واحد منها فقط أقل من ٠,٨٠ وكذلك كانت معاملات الاستقرار بإعادة الاختبار بعد فترة أسبوعين وكذلك بعد فترة ثلاثين يوما عالية أيضاً، كما أثبتت دراسات طولية بدرجة استقرار مناسبة لفترة طويلة للدرجات على المقاييس المفردة، وكانت هذه الدراسات تمتد بين ثلاث سنوات وثمان سنوات، وامتد بعضها إلى ٣٠ سنة أو أكثر (٥٨).

أما صدق اختبار سترونج، فيوجد قدر كاف من نتائج البحوث يبين أن

المفحوصين يميلون إلى الالتحاق والاستمرار فى المهن التى حصلوا فيها على درجات مرتفعة على الاختبار.

وبالنسبة إلى تلاميذ الثانوى فإنه من الممكن التنبؤ بدرجة حسنة بالمجال أو جماعة المهنة التى سيلتحق بها التلميذ، أما التنبؤ بمهنة بعينها فإنه لايوثق به تماما إلا فى حالة تلك القلة من التلاميذ التى تظهر عندها أنماط محددة ومميزة جدا من الميول.

ويتوافر قدر من البيانات التى توضح وجود ارتباط بين درجات اختبار سترونج والنجاح فى المهنة أو الرضا عنها، أما الأدلة التى تثبت عكس هذا فهى نادرة ويصعب تفسيرها.

والواقع أن المزايا التى يتمتع بها اختبار سترونج جعلت أنستازى تذهب فى تقديره إلى القول بأنه «واحد من أنجح الطرق لقياس المتغيرات غير العقلية» (١٨ : ٥٣٦)، وما زال الاختبار خاضعا لعدد من مشروعات البحث فى «مركز البحث فى قياس الميول» The Center for Interest Measurements Research الذى أسس فى عام ١٩٦٣ فى جامعة مينسوتا بالولايات المتحدة، ويكفى أن نشير إلى أن واحدا من هذه المشروعات يتضمن دراسة تمتد ٣٥ عاما لأكثر من ألف تلميذ عمرهم ١٥ سنة، ومن المتوقع أن تزيد هذه الدراسات من قيمة الاختبار، وقد نقل هذا المقياس إلى اللغة العربية عطية محمود هنا واستخدم فى عدد من البحوث المصرية والعربية.

وقد تعرض اختبار الميول المهنية لتسرونج إلى تعديلات جوهرية بدأت منذ عام ١٩٧٠ واصدرت العديد من التعديلات. وتحدد اناستازى واريانا (١٩٩٧- ٣٩٠) التعديلات الرئيسية:

١- وجود اطار نظرى مرجعى للمؤسسات التى ستستخدمه وكيفية تفسير الدرجات.

٢- بداية لاستحداث مقياس للنساء وآخر للرجال مع وضع معايير جديدة لكل مقاييس المهن القديمة الخاصة بالنساء والرجال، كما استخدمت عينات جديدة.

٣- زيادة عدد المقاييس المهنية والفنية للمهن التى تتطلب درجات علمية أقل

من التعليم الجامعى. ويعد هذا تعديلا جوهريا لعدم تضمينه فى الطبعات السابقة.

وصف اختبار الميول المهنية المعدل (SII -FORM T317) يتألف اختبار الميول المهنية المعدل من ٣١٧ مفردة تتجمع فى ٨ أقسام، ويوجد فى الاقسام الخمسة الأولى استجابات المفحوص لتفضيلاته وذلك بوضع الحرف "L" أو "I" أو "D" والتى ترمز الى (Like -Indifferent - Dislike) وتقع المفردات فى الاقسام الخمسة الأولى فى الفئات الاتية: المهن – المواد الدراسية – الأنشطة (مثل الخطابة – اصلاح الساعات – التبرع بالمال) أنشطة وقت الفراغ – الاختلاط مع الناس من فئات مختلفة (مثل كبار السن) أو رجال القانون .

وفى القسمين التاليين يطلب من المفحوص التعبير عن تفضيله لازواج من الانشطة مثل (التعامل مع الاشياء – التعامل مع الناس) وبين كل الثنائيات المحتملة لاربعة مفردات ترتبط بالعمل (افكار – تواريخ – اشياء – افراد) وفى القسم الاخير يطلب من المفحوص ان يحدد قائمة من الصفات الشخصية بوضع نعم – لا.

تصحيح الاختبار:

يصحح الاختبار باستخدام الحاسب الآلى فى المراكز المتخصصة والتى حددها ناشر الاختبار أوالحصول على C-D من ناشر الاختبار. ويوجد ثلاث مستويات للتصحيح تختلف فى نوع ودرجة شيوع المهنة وقد حدد ستة مهن رئسية فى المستوى الأول، وتضمن المستوى الثانى ٢٥ اختبارا فى الميول الاساسية، والمستوى الثالث يتضمن ٢١١ مقياسا للميول المهنية وبالاضافة الى (اختبار ٣١٧) (T.317) يقدم اختبار سترونج درجات لأربع اساليب شخصية والتى تقدر تفضيل المفحوص لاسلوب العمل، بيئة التعلم ، نمط القيادة والمخاطرة ويقدم نموذج التصحيح (البروفيل) ايضا مجموعة من المؤشرات التطبيقية والتى من خلالها يمكن التعرف على الاهمال فى الاستجابة أو اسلوب الاستجابة Response set للمفحوص.

ويعتمد اختبار سترونج للميول المهنية على تنظير طوره جون هولاند (١٩٦٦، ١٩٨٥، ١٩٩٢) والذى دعم بالعديد من البحوث من هولاند نفسه وآخرون.

والفئات المهنية التى عرضها هولاند فى برنامجه هى: الواقعى Realistic (R)، المقتصى Investigative (I)، الفنى Artistic (A)، الاجتماعى Social (S)، المغامر

Enterprising (E)، التقليدى Conventional (C)وكل من هذه الفئات تعبر عن نمط الشخص وأيضا بيئة العمل المناسبة. ويحدد هولاند ان هذا التصنيف ليس جامدا فى الفئات الستة.

وتصنف الاختبارات الخمسة والعشرين الاساسية للفئات الستة للمهن العامة وتتألف هذه الاختبارات من تجمعات Clusters بين مفرداتها ارتباطات Inter-correlated، واختبارالميول الاساسى مفرداته أكثر تماسكا من اختبارات المهن وعلى ذلك يمكن تفسير لماذا يحصل فرد ما على درجات مرتفعة فى اختبار ما للمهن.

وقد استخدمت العديد من البحوث هذا الاختبار لتحديد صدقه وثباته فبالنسبة لثبات اختبار الميول المهنية كان وسيط معاملات الثبات ٠,٩ من عينة قوامها ١٩١ من العمال المعينين وكان الفاصل الزمنى بين التطبيقيين تراوح بين ثلاث وستة اشهر.

اما صدق الاختبار فقد تم تحيده بمحكات تلازمية وايضا تنبؤية على عينات مختلفة فى مهن مختلفة وعينات مرجعية.

ثانيا : اختبار كيودر للميول المهنية:

Kuder Preference Record-Vocational

يعتبر هذا الاختبار أكثر مجموعة اختبارات كيودر للميول شهرة أو استعمالا، والواقع أن الصور المختلفة، والطبعات المعادة والمعدلة لاختبارات كيودر للميول تمثل عائلة من الأدوات التى تتخذ لقياس الميول سبلا وزوايا مختلفة، كما أنها مصممة لأغراض مختلفة، إلا أن أهمها هو هذا الذى نتكلم عنه، وقد وضع هذا الاختبار كيودر، واتبع منهجا مختلفا، بل مضادا، لمنهج سترونج فى اختيار وتصحيح عناصر الاختبار، لقد كان هدفه الرئيسى هو بيان الميل النسبى فى عدد صغير من المجالات الواسعة، وليس فى مهنة بذاتها، وقد صيغت عباراته أصلا وجمعت مبدئيا على أساس صدق المحتوى، وقد اتبع هذا تحليل شامل لعناصر الاختبار من تطبيقها على جماعات من تلاميذ الثانوى والكبار، وكان الهدف من تحليل العناصر هذا هو إنشاء مجموعات عناصر ذات اتساق داخلى عال وارتباط منخفض بالمجموعات الأخرى وقد تحقق هذا الهدف بشكل جيد فى معظم مقاييس الاختبار.

وعناصر اختبار كيودر من نوع الاختيار الاجبارى الثلاثى، وقد تعرفنا عليه فى بعض اختبارات الشخصية من قبل - وعلى المفحوص أن يختار بين الأنشطة الثلاثة المذكورة ذلك النشاط الذى يحبه أكثر، كما يبين النشاط الذى يلقى عنده أقل تفضيل، ومن أمثلة هذه الثلاثيات ما يلى، وهى مختارة من الصورة العربية للاختبار التى أعدها الدكتور أحمد زكى صالح (ص ٧ من كراسة الأسئلة).

١٦ – تعمل نموذجا من الصلصال.
١٧ – تكتب مقالا عن طريقة إقناع الناس وتوجيههم.
١٨ – تعمل ملقنا فى تمثيلية يقوم بها هواه.
٩ – تحترف الطب.
٢٠ – تحترف النحت.
٢١ – تحترف الصحافة.

يعطى الاختبار عشرة مقاييس ميل بالإضافة إلى مقياس تحقق Verification scale للكشف عن الإهمال، أو سوء الفهم، واختيار الإجابات المقبولة والمرغوب فيها اجتماعيا ولكنها غير متوقعة، وتتضمن مقاييس الميول العشرة الميول الآتية: الخلوى، والميكانيكى، والحسابى، والعلمى، والاقناعى، والفنى، والأدبى، والموسيقى، والعدالة الاجتماعية، والكتابى، وقد وضع كيودر للاختبار معايير خاصة بكل من الذكور والإناث لتلاميذ الثانوى، والجامعة، والكبار، وتخطط الدرجات الكلية على مجالات الميول العشرة على قائمة معايير مئينية.

وتدور معاملات ثبات اختبار كيودر حول ٠,٩٠، كما يبدو الثبات بإعادة الاختبار بعد عام أو أقل مرضيا، أما بعد فترات أطول فليس هناك قدر كاف من المعلومات عنها، وما هو موجود يشير إلى أنه، وخاصة حالة تلاميذ الثانوى، كانت هناك تغيرات فى مجالات الميول المرتفعة والمنخفضة عندما كانت الفترة بين الاختبارات تمتد إلى عدد من السنين (١٨ : ٤٧٥) ويبدو أن كلا من اختبار سترونج وكيودر عرضة للكذب أو التحريف، ويزيد هذا بالنسبة إلى اختبار كيودر نظرا لطبيعة عناصره الأكثر وضوحا وشفافية.

أما دراسات صدق الاختبار فقد أجريت أساسا بمقارنة الاختبار بمحك الرضا فى

العمل، ومن أكبر هذه الدراسات الطولية الشاملة تلك التى أجريت على ١١٦٤ تلميذا أعطوا اختبار كيودر فى الثانوى، ثم أعطوا استبيانا للرضا فى العمل بعد عشر سنوات، وعند ذاك كان عدد ٧٢٨ من هؤلاء ملتحقين بأعمال «متسقة» مع أنماط ميولهم الأصلية، و٤٣٦ فى أعمال «غير متسقة» مع هذه الأنماط وكانت نسبة العاملين الراضين فى أعمالهم فى الجماعة المتسقة هى ٦٢٪، بينما كانت النسبة هى ٣٤٪ فى الجماعة غير المتسقة، ومن ناحية أخرى كانت نسبة العاملين غير الراضين ٨٪ فى الجماعة المتسقة، و٢٥٪ فى الجماعة غير المتسقة (١٨)، وقد أيدت هذه النتائج دراسات أخرى فى ذات الاتجاه.

وقد تطور حديثا عن اختبار كيودر اختبار آخر هو

اختبار كيودر لمسح الميل العام (٤٠) Kuder General Interest Survey.

وهو عبارة عن مراجعة وامتداد للاختبار الأصلى فى الاتجاه الأدنى من الميول وقد صمم للأعمار من نحو ١٢ حتى ٢٤، ويستعمل لغة أبسط وألفاظ أسهل من الاختبار السابق، ولهذا فهو لايتطلب مستوى رفيعا من القراءة، وقد وضعت معايير مئينية أمريكية لأكثر من عشرة آلاف ولد وبنت، ولعينات قليلة من الكبار ومازال البحث جاريا للتحقق مما إذا كانت هذه الأداة تناسب الكبار كما تناسب من هم أصغر سنا، وفى هذه الحالة قد ينتهى الأمر إلى أن يحل اختبار كيودر لمسح الميل العام محل الاختبار السابق، وتشير الدراسات التى أجريت على الاختبارين أن الارتباطات بينهما تعادل معاملات الثبات الخاصة بهما، أى أنها تشير إلى درجة التشابه الكبير فى قياس الميول عند الكبار.

وثمة صورة أخرى لاختباركيودر للميول المهنية وهى اختبار كيودر لمسح الميل المهنى(٤١). Kuder Occupational Intrest Survey (KOIS)

وقد تطور هذا الاختبار وأنشئ على أساس المحك وبطريقة متشابهة لتلك التى اتبعها ستورنج فى اختباره، كما سبق بيانه، إلا أنه فى هذه الصورة لم يستعمل جماعة الرجال بصفة عامة كما هو الحال عند سترونج، وبدلا منها فإن درجة المفحوص على كل مقياس مهنة يعبر عنها ارتباط بين نمط ميوله ونمط ميول جماعة مهنية معينة، ويمكن الحصول حاليا على درجات لعدد ٧٩ مهنة، و٢٠ مجالا

التقويم النفسى ـــــــ

باهتمام علمى واسع فى مصر، واستعمل ولايزال فى عدد من البحوث فى مجال الميول المهنية (١ : ٢٦١ – ٣١٤).

ثالثا : اختبار لى - ثورب The Lee - Thorpe Inventory

بينما يمثل اختبار سترونج الأساس التجريبى فى بنائه وتصحيحه، ويمثل كيودر التنوع والتعدد، نجد أن اختبار لى - ثورب يمثل اتجاها ثالثا فى بناء اختبارات الميول، هذا الاختبار يتكون من مجموعة من الأمثلة التى اختيرت ونظمت على أساس الحكم وليس على الأساس الإحصائى، هذا الاتجاه الذى يمكن أن يسمى الاتجاه المنطقى، مضاد لاتجاه سترونج الذى يعتمد على النتائج الإحصائية، ويشبه هذا الاتجاه المنطقى الأسلوب الذى يتبع فى بناء اختبارات القدرة والاستعداد حيث تعين مجموعة كبيرة جداً من المواقف ثم تختار عنصر عشوائيا من هذه المجموعة الضخمة.

وقد بدأ لى ثورب بوصف المهن الموجودة فى قاموس العناوين المهنية Dictionary of Occupational Titles الذى يصف ويصنف كل المهن والأعمال تقريبا فى المجتمع الأمريكى ، واختيرت فى ستة مجالات مهام تمثل ثلاثة مستويات للمسئولية: مرتفع ، ومتوسط، ومنخفض وتذكر أوصاف موجزة للأعمال فى أزواج، ويبين المفحوص ما يفضله من واحد من زوج، وتمثل الدرجة التكرار النسبى للاختيارات في أقسام: الشخصية الاجتماعية،والطبيعية، والميكانيكية، والأعمال، والفنون، والعلوم.

ويختلف هذا الاختبار عن اختبار سترونج فى أن الاختبار الأصلى وتصنيف عناصره يعتمد كليه على أوصاف للمهنة وليس على دليل واقعى على أن أشخاصاً فى المهنة يحبون نشاطا معينا كما يختلف عن اختبار كيودر فى أن التجميع الأول للعناصر بنى على أساس تحليل منطقى وليس على أساس العزل الإحصائى للعوامل وعلى الرغم من أن التصنيف الأصلى للعناصر أعتمد كليا على أساس منطقى فقد أجريت دراسات ارتباطية تالية بهدف رفع مستوى التجانس، وعند مراجعة الاختبار استبعدت العناصر التى لاتشترك مع بقية العناصر فى القسم الواحد (٢٧).

الميول والشخصية : اتجاهات وتوقعات :

كانت الدراسات الأولى فى الميول تعتمد اعتمادا كبيرا على المقارنة الواقعية

الخالصة لجماعات مهنية على أساس عناصر مختارة بطريقة عشوائية غالبا، أى أنه لم يكن هناك نظرية خاصة بطبيعة الميول، ولابأنماط الميول التى يجب أن نوليها عناية خاصة، وكذلك لم تكن هناك نظرية عن تكوين المهن والأعمال، وكانت تفسيرات درجات اختبارات الميول تتم على أساس عملى غير نفسى (٢٧ : ٤٣٦)، وقد تغير الاتجاه فى قياس الميول منذ بداية نشر سترونج اختباره، ويمكن القول أن اختبار سترونج واختبار كيودر أديا إلى بداية نظرية فى الميول، وقد وجد جيلفورد وزملاؤه (٣١) أكثر من عشرين عامل ميل يبدو أن معظمها يعكس أساليب شخصية عامة وليس اتجاهات مهنية، وأضاف جيلفورد إلى ما هو معروف من ميول مثل الميكانيكية والعلمية وغيرها، وأضاف المغامرة فى مقابل الأمن، والتذوق الجمالى، والمسايرة الثقافية أو النظام، والعدوان وأبعاد ميول أخرى كثيرة وكذلك أدت الدراسات الطولية ودراسات الحالة إلى إعطاء صورة أكثر وضوحا عن أهمية الميول ومغزاها، أى أن هناك تطورا ونموا متوقعا فى اتجاه نظرية متكاملة عن الميول.

وثمة اتجاه آخر نام أيضاً هو التفسير المنتظم للميول فى ضوء مفاهيم أساسية للشخصية، فبينما أتى على الميول حين من الدهر كان ينظر إليها فيه على أنها عمليات اشتراك تقوم على الصدفة، نجد أن الميول الآن ينظر إليها على أنها تعبير عن حاجات عميقة، وعن أنماط توافق معينة، وربما تكون اختبارات الميول أكثر كفاءة من عدد كبير من الأساليب الأخرى فى قياس الشخصية نظرا لتنوع محتواها وزيادة تقبل الشخص بها.

وقد كشف عدد كبير من الدراسات عن ارتباطات ذات دلالة بين الميول المهنية المقاسة وجوانب أخرى للشخصية، فمن هذه الدراسات ما أظهر مثلا أنه فى حالات إدمان الخمر، والجنسية المثلية، والعصاب، وغيرها من نواحى العجز الذهانية كانت تسود كل جماعة من هذه الجماعات أنماط ميول مهنية خاصة مميزة لها (٢٥، ٢٨)، وبينت دراسات أخرى أجريت على تلاميذ المدارس والجامعات والعمال والصناعيين والمرضى الراشدين أن الاستهداف للحوادث يرتبط ارتباطا ذا دلالة بمقاييس معينة على اختبار سترونج، وكذلك ثبت أن هناك ارتباطا ذا دلالة بمقاييس معينة على اختبار سترونج، وكذلك ثبت أن هناك ارتباطا بين درجات اختبار سترونج وكيودر واختبارات الشخصية مثل اختبار مينسوتا المتعدد الأوجه وغيره (١٨ : ٤٩١).

ومن جهة أخرى فإنه من المعترف به الآن أن اختبار مهنة غالبا ما يعكس حاجات الشخص العاطفية الأساسية، وأن التوافق المهنى يمثل جانبا أساسيا فى توافق الشخص العام فى الحياة، ذلك لأن الشخص عندما يختار مهنة، فإنه إلى حد بعيد، يختار أساليب التوافق وأنماط الحياة والأدوار التى يحس أنها قريبة إليه ومناسبة له، وعلى هذا فإن قياس الميول المهنية، وبصفة خاصة تحديد الجماعات المهنية التى يشارك افراد أعضاءها ميولهم واتجاهاتهم، يصبح أمرا هاما فى فهم مختلف الشخصيات، ومما يؤيد التأكيد الحالى لاتخاذ الاختبارات المهنية مؤشرات للشخصية ذلك الاختبار الذى وضعته هولاند (٣٠) Holland Vocational Preference In-ventory وقد استعمل فيه العناوين المهنية لقياس خصائص الشخصية، وما على المفحوص فى هذا الاختبار إلا أن يوضح ما إذا كان يحب أو يكره كل واحدة من ثلاثمائة مهنة، وعلى الرغم من أن هذا الاختبار قد اعتمد اعتمادا كبيرا على صدق المحتوى الداخلى، فإن مقاييسه قد أظهرت اختلافات دالة بين عينتين متناظرتين أحدهما تتكون من أفراد مصابين بأمراض نفسية، وتتكون الأخرى من أفراد عاديين ويعطى الاختبار عشرة مقاييس للشخصية، ومنها النشاط البدنى والنزعة العقلية، والمسئولية، والمسايرة ومازالت قيمة هذا الاختبار رهن البحث والاستعمال العملى إلا أنه يمثل اتجاها جديدا فى دراسة الشخصية عن طريق الميول.

من هذا يتضح أنه من المتوقع أن تتطور نظرية خاصة بالميل ، وأن يمثل قياس الميول مجالا حيا فى تطور الاختبارات وفى البحث العلمى، وأن تنمو وتتنوع الدراسات حول العلاقات بين الميول الشخصية ، وأن يتجه تفسير اختبارات الميول إلى مزيد من العمق والدقة .

ومن الجهود المصرية فى اعداد مقاييس للميول مقياس الميول نحو المواد الدراسى والذى اعده فؤاد ابو حطب للبيئة المصرية والعربية واستخدم فى عدد من البحوث. وفيما يلى عرض للمقياس :

مقياس الميول نحو المواد الدراسية

وهو من اعداد فؤاد ابوحطب عن اصل للباحثين الانجليزيين Duckworth, D.Z Entwistle N.S. دراستهما المنشورة فى عام ١٩٧٤ (١٧ : ٧٦-٨٣) والتى

افترضت ان هذا المقياس يتناول فى جوهره الاتجاهات نحو المواد الدراسية الا ان فؤاد ابو حطب (٢) قام بتعديل هذا المقياس وطور فى اطاره النظرى بحيث جعله اكثر ملائمة لقياس الميول Interestsمنه إلى قياس الاتجاهات Attitudes ويعتمد هذا المقياس على نظرية جورج كيلى المعرفية فى الشخصية وطريقته فى قياسها التى يسميها الاسلوب الشبكى للمحصول Repertory Gridtechnique والتى وصفها بانستر بالتفصيل (٥) وكانت الخطوة الاولى والاساسية فى بناء المقياس هى تعيين التكوينات الملائمة بشرط ان تكون هذه التكوينات ذا معنى لدى الافراد موضوع القياس.

وكانت الطريقة الاصلية التى لجأ اليها جورج كيلى هى استحضار هذه التكوينات عن طريق المقارنات المتكررة بين مجموعات مختلفة من الاشياء كما انه استخدم طريقة المناقشة، الا ان الطريقة التى لجأ اليها دكوورث وانتوستل هى اجراء مقابلات شخصية مع تلاميذ المرحلة الثانوية الذين سئلوا ان يفسروا لماذا اختاروا أو رفضوا مواد دراسية معينة، وانشئت عبارات مختلفة حول الاسباب المقترحة ثم خضع المقياس للتحليل العاملى وتوصل الباحثان الى عوامل اربعة هى Interestوالحرية feeedom والصعوبة difficulty والمنفعة social benefitوجاء المقياس فى صورته النهائية يتألف من ١٨ عبارة خمس عبارات لكل بعد من ابعاده فيما عدا بعد المنفعة فكان عدد عبارات ثلاثة فقط.

وقد قام فؤاد ابو حطب باعادة بناء هذا المقياس فأصبح فى صورته العربية يتألف من ٢٠ عبارة لكل بعد عدد متساو من العبارات (٥ عبارات) وقام بتقنينه على عينات مختلفة من تلاميذ المرحلتين المتوسطة والثانوية بالاضافة الى مرحلة التعليم الجامعى فى المملكة العربية السعودية وفى هذا التعديل للمقياس اقترح تسميته مقياس الميول نحو المواد الدراسية والابعاد الاربعة فى رأى فؤاد ابو حطب تنتمى الى بعد الميول الذى يغلب عليه الجانب الوجدانى، ولا تتجاوزه الى الجانب المعرفى كما هو الحال فى الاتجاهات.

ويتكون المقياس من قائمتين من الصفات التى يمكن ان توصف بها المواد الدراسية هما القائمة (أ) والقائمة (ب) وبينهما يوجد عمود يدل على مادة دراسية

معينة، والمطلوب من المفحوصين أن يقرأ كل صفة فى القائمة (أ) وما يقابلها فى القائمة (ب) ثم يختار من بينهما الصفة التى يجدها انسب لوصف رأيه فى هذه المادة الدراسية وذلك بأن يضع تحت اسم المادة الرمز (أ) إذا اختار لها الصفة فى القائمة (أ) أو الرمز (ب) إذا اختار الصفة الموجودة فى القائمة (ب).

ويوجد بكل من القائمتين (أ) و(ب) ٢٠ صفة متقابلة أو عشرون بندا يضاد كل بند البند المقابل له.

وتعبر بنود المقياس عن الأبعاد الآتية:

١- الاهتمام.

٢- الحرية

٣- الصعوبة

٤- المنفعة .

ومثال لاحد المفردات فى المقياس.

كثيرا ما تجعلنى أشعر بالسأم (مقابل) عادة ما تثير اهتمامى وينتمى هذان البندان الى بعد الاهتمام.

ويتضمن المقياس مفتاحا لتصحيح استجابات المفحوص حيث يحصل على درجة واحدة إذا انطبقت استجابة المفحوص ومفتاح التصحيح.

الخصائص السيكومترية للمقياس:

تم اجراء صدق المقياس فى دراسة راجية شكرى سنة ١٩٨١ باستخدام محك تقديرات المعلمين، فقد طلب من كل معلم أن يحدد أكثر خمسة تلاميذ يقوم بتدريسهم يرى أنهم يميلون الى المادة الدراسية وكذلك اختيار أقل خمسة تلاميذ يظهرون ميلا نحو المادة الدراسية.

وقد تم مقارنة اداء المجموعتين باستخدام اختبار «ت» التى تراوحت قيمتها بين ٧,٧٣ - ١٣,٦٣ لصالح المجموعة الاكثر ميلا نحو المادة الدراسية فكانت كل القيم دالة عند مستوى ٠,٠١ .

وفى دراسة شحاته محروس طه سنة ١٩٨٥ تم حساب صدق الاختبار باستخدام المجموعات المتضادة كمحك لمقارنة اداء طلاب كلية التربية الموسيقية حيث تم مقارنة اداء الارباعى الاعلى/ الارباعى الادنى فى المواد الدراسية وتم مقارنة اداءهم فى مقياس الميول فى المواد الدراسية وكانت قيم ت جميعها دالة عند مستوى ٠,٠١ لصالح المجموعة الاعلى تحصيلا فى المقررات النظرية، اما بالنسبة للمقررات العملية فكانت جميعها دالة ايضا فيما عدا بعد الحرية وربما يفسر ذلك بأن الاداء الموسيقى للأعمال الموسيقية محدد ولا يستطيع الطالب الخروج عما هو مطلوب من المؤلف.

أما ثبات المقياس فقد تم تحديده فى الدراستين بإعادة التطبيق وتراوحت المعاملات فى الدراسة الاولى بين ٠,٣٤ ، ٠,٧١ وفى الدراسة الثانية بين ٠,٧٥٣ ، ٠,٩٢٤ وكانت جميع المعاملات دالة.

مراجع الباب السادس

١-أ: آمال أحمد مختار صادق: تقنين اختبار البروفيل الشخصى لجوردون فى: بحوث فى تقنين الاختبارات النفسية تحرير فؤاد أبو حطب – المجلد الأول الانجلو المصرية ١٩٧٧ .

١ – ب: ـــــــــــــ : تقنين اختبار قائمة الشخصية لجوردون: فى بحوث فى تقنين الاختبارات النفسية تحرير فؤاد ابو حطب– المجلد الاول – الانجلو المصرية ١٩٧٧ .

١–جـ: ـــــــــــــ : اختبار الاتجاهات التربوية – فى بحوث فى تقنين الاختبارات النفسية تحرير فؤاد ابو حطب– المجلد الثانى – الانجلو المصرية ١٩٧٩ وكراسة التعليمات .

١ – أحمد زكى صالح: علم النفس التربوى ، القاهرة مكتبة النهضة المصرية (ط ١٠) ، ١٩٧٢ .

٢ – أحمد زكى صالح: الأسس النفسية للتعليم الثانوى ، القاهرة مكتبة النهضة المصرية (ط ٢) ، ١٩٧٢ .

٣– أمال عبد السميع باظة: مقياس السلوك العدوانى للاطفال – كراسة التعليمات – الانجلو المصرية ٢٠٠٠ .

٣أ – بانسترب : نظرية جديدة فى الشخصية، الفصل التاسع عشر من كتاب آفاق جديدة فى علم النفس تحرير ب .م فوس – ترجمة فؤاد ابو حطب عالم الكتب ١٩٧٢ .

٣ ب– ـــــــ اختبار الأسلوب المعرفى – كراسة التعليمات – الانجلو المصرية ٢٠٠٠ .

٣ جـ: اسماء السرسى وامانى عبد المقصود – الكفاءة الاجتماعية لاطفال ما قبل المدرسة الابتدائية– كراسة التعليمات – الانجلو المصرية ١٩٨١ .

٣ د– ايمان فؤاد كاشف وآخر : المقياس الموضوعى السلوكى لذوى الاعاقة العقلية – كراسة التعليمات – الانجلو المصرية ٢٠٠٥ .

٣ هـ– خالد فؤاد ابو حطب : مقياس كامدكس لقياس خرف الشيخوخة، تحت الطبع.

٣ و- راجية محمد شكرى: أثر اسلوب التدريس فى تحصيل التلاميذ وميولهم نحو المادة الدراسية رسالة ماجستير غير منشورة - كلية التربية - جامعة عين شمس ١٩٨١ .

٤ - أحمد عبد العزيز سلام: تطبيق اختبار تفهم الموضوع على حالات مصرية، رسالة ماجستير، كلية التربية جامعة عين شمس، ١٩٦٢ .

٤ أ- سيد أحمد عثمان: المسئولية الاجتماعية، القاهرة مكتبة الأنجلو المصرية، ١٩٧٣ .

٥ - سيد أحمد عثمان: علم النفس الاجتماعى التربوى، القاهرة مكتبة الأنجلو المصرية، ١٩٧٠ .

٦ - سيد أحمد عثمان، فؤاد أبو حطب: التفكير - دراسات نفسية، القاهرة: مكتبة الأنجلو المصرية، ط (٢) ١٩٧٧ .

٧ - سيد محمد غنيم: مدى صلاحية اختبار بقع الحبر لرورشاخ لقياس الذكاء، رسالة ماجستير، كلية التربية جامعة عين شمس، ١٩٥٥ .

٨ - سيد محمد غنيم، هدى عبدالحميد برادة: التشخيص النفسى، دراسات فى اختبار الرورشاخ، القاهرة: دار النهضة العربية (جـ ١)، ١٩٦٥، (جـ ٢) ١٩٦٦ .

٩ - سيد محمد غنيم، هدى عبدالحميد برادة: الاختبارات الإسقاطية، دراسات فى اختبار الرورشاخ، القاهرة: دار النهضة العربية ١٩٦٤ .

١٠ - فاروق إبراهيم أبو عوف: دراسة مقارنة لمعرفة مدى الإفادة من اختبار رورشاخ (بقع الحبر) للتمييز بين حالات العصاب القهرى والأسوياء، رسالة ماجستير، كلية التربية جامعة عين شمس، ١٩٧٢ .

١١ - فاروق إبراهيم أبو عوف: مدى صلاحية اختبار رورشاخ لقياس الابتكارية، رسالة دكتوراه، كلية التربية جامعة عين شمس، ١٩٧٦ .

١٢ - فاروق إبراهيم أبو عوف: دراسة مقارنة لبعض سمات الشخصية بين القادة وغيرالقادة باستخدام اختبار بقع الحبر لرورشاخ، الكتاب السنوى فى علم النفس، الجمعية المصرية للدراسات النفسية، المجلد الرابع، القاهرة، مكتبة الأنجلو المصرية، ١٩٨٥ .

١٣ - فؤاد أبو حطب: التحليل العلمى للسلوك الخلقى، الكتاب السنوى فى التربية وعلم

النفس (تحرير سعيد إسماعيل على) ، المجلد الأول، القاهرة: عالم الكتب، ١٩٧٣ .
١٣أ- فؤاد أبو حطب: التفضيل الفنى وسمات الشخصية، المجلة الاجتماعية القومية، يناير ١٩٧٣ .
١٤ أ- فؤاد أبو حطب : اعداد وتقنين مقياس الميول نحو المواد الدراسية - الانجلو المصرية ١٩٩١ .
١٤ب - ـــــــــ اعداد وتقنين اختبار البروفيل الشخصى لجوردون دار النهضة العربية ١٩٧٢ .
١٤جـ- ـــــــــ اعداد وتقنين اختبار قائمة الشخصية لجوردون ، دار النهضة العربية ١٩٧٢ .
١٤ د: فاروق عبد الفتاح موسي وآخر: مقياس مفهوم الذات للاطفال - كراسة التعليمات - مكتبة النهضة المصرية ١٩٩٦ .
١٤ هـ- فوقيه عبد الفتاح: بناء وتقنين مقياس التفكير الاخلاقى للاطفال - كراسة التعليمات - الانجلو المصرية ٢٠٠١ .
١٥ - لويس كامل مليكه، محمد عماد الدين إسماعيل: الشخصية وقياسها، مكتبة النهصة المصرية، ١٩٥٩ .
١٦ - نجيب اسكندر إبراهيم، لويس كامل مليكه، رشدى فام منصور: الدراسة العلمية للسلوك الاجتماعى، القاهرة: مؤسسة المطبوعات الحديثة، ١٩٦٦ .
١٧ - يوسف الشيخ، جابر عبد الحميد جابر: سيكولوجية الفروق الفردية، القاهرة: دار النهضة العربية، ١٩٦٤ .

17.A. Abu Hatab, Maha: Comon Behavioral Prolens Among Sample of Egyptian Preschool Children. A Teacher Rating Perspective J. Current Paychiatry Vol 10, No2, 2003 .

17.B. ---------- Childhood Phobia: Screening Among a Sample of non - Referred Primary School Egyptian children. Current Psychiatry Vol. II No1, 2004 .

17.C. ---------- School Refusal : A Community Sample Assessment J. Current Psychiatry, Vol. 12 No1, 2005 .

18. Anastasi A. Psychological Testing. NewYork, Macmillan, 1982.

18 A. Anastasi A & Urbina, S., Psychological Testing. NewYork,

Macmillan, 1997.

19. . Atkinson, J.W. (ed.) Motivation, Fantasy Action, and Society, Princeton, N.J. Van Nosfland, 1958.

20. Atkinson, J.W. and Feather, N.T. A theory of Achievement Motivation. New York : Wiley, 1966.

20A. Barker, F.B. Computer Techonology in Test ConStruction and Processing In : Lin R.L. ed: Educational Measurement, New York, A Merician Council in Education 1989 .

21. Barkes, E.G., Kounin, G. and Wright, H.F. (eds.) Child Behaviour and Development. New York : McGraw-Hill, 1943.

22. Beck, S.L. Rorchach's test. I : Basic Processes. New York : Grune and Stratton, 1944.

23. Beck, S.L. Rorchach's test. Ill : Advances in iInterpretation New York : Grune and Stratton, 1952.

24. Bell, H.M. The Adjustment Inventory. Palto Alto, Calif: Consulting Psychologists, 1934-1939.

24A. Butcher, J. N., Computerized Psychological Assessment, New York Basic Books, 1987 .

25. Carnes, G.D. Vocational Interest Characteristic of Abnormal Personalities. J. of Counsel, Psychol. 1964, pp, 272-279.

26. Combs. A.W., Snygg, D. Individual Behaviour, New York : Harper, 1959.

27. Cronbach, G.L. Essentials of Psychological Testing, New York : Harper, 1984.

28. Drasgow, J. and Carkhuff, R.R. Kuder Neuro-Psychiatric keys before and after psychotherapy. J. of Counsel. Psychol. 1964,11,67-71.

28A. Duckworth, D.Z., Entwistle, Attitudes Towards School Subjects: A Repertory Grid Texhnique British J.Educ. Psy. vol 44, 1974 , pp

76-83.

29. Flanagan, G.c. Factor Analysis In The Study Of Personality .stanford : Stanford Univ. Press, :1935.

30 . goldner, R,h. Individual Differences In Whole-part Approach And Flexibility-rigidity In Problem Solving. Psycholo. Monogr. 1957, 71, No. 21.

31. Guilford, L.p. Et Al., A Factor Analysis Study Of Human Interests. Psychol. Morograph, 1954, 68, No. 41.

31 A. Gutkin, T. B., & Wise, S., The Computer And Decision Making Process. Hills - Dale, N.j: Erlbaum 1991 .

32. Hartshone; H. And May, Studies In Deceit. New York : Macmillan, 1928.

33. Hartshome, H. And.may, M. Studies In Service And Self- Control, New York : Macmillan, 1929.

34. Hartshorme, H., & May, M., Studies In The Organization Of Character, New York, Macmillion 1930 .

35. Hathaway, S.r., And Mckinley, G.c. Minnesota Mackinley, G.c. Personality Inventory : Manual For Administration And Scoring. New York : Psychological Corporation, 1967.

36. Holland, J.l. A Personality Inventory Employing Occupational Titles. J. Of Appl. Psychol., 1958, 42, 336-342.

37. Holtzmann, W.h., Et Al. Inkblot Perception And Personality-Holtzmann Inkblot Technique. Austin, Texas : Univ. Of Texas Press, 1961.

37a. Jensen R., Arthur, Bias In Mental Testing, Methuen And Co, U.s.a. 1980 .

38. Klein, G.s. And Schlesinger, H.l, Perceptual Attitudes Toward Instability : I. Prediction Of Apparent Movement Experiences From Research Response. J. Personol., 1951, 19, 289-302.

39. Krech, D., Crutchfield, R.s. And Ballachey, E. L. Individual In Socicty. New York :mcgraw-hill, 1962.

40. Kuder, G.f. General Interest Survey, Manual Chicago : Science Research Associates, 1964.

41. Kuder, G.f. Kuder Occupational Interest Survey : General Manual. Chicago ; Science-research Associates, 1966.

42. Kunce, J.t. Vocational Interests And Accident Proneness. J. Of Appl. Psychol, 1967, 51, 223.

43. Lindzey, G. And Borgatta, E.f. Sociometric Measurement. In Lindzey, G. (Ed.) Handbook Of Social Psychology. Cambridge, Mass. Addison-wesley, 1954, Vol. I, Ch. Ii.

44. Luchins, A.s. Mechanization In Problem Solving. Psychol. Mon-og:, 1942, 54, No. 6.

45. Masling, J. The Influence Of Situational And Interpersonal Variables In Projective Testing. Psychol. Bulletin, 1960, 56, 65-85.

46. Masling, J. Differential Introduction Of Examiners And Rorcharch Responses. J. Of Consult. Psychol., 1965, 29, 198-201.

47. Mcclelland, D.c., Et Al. The Achievement Motive. New York: Appieton-century, 1953.

48. Mcclelland, D.c. The Achieving Society. New York : Free Press, 1961.

49. Moreno, G.l. Who Shall Survive ? Washington : Nervous And Mental Diseases Publishing Co., 1934.

50. Murray, H. A., Et Al. Explorations In Personality. New York: Oxford Univ. Press, 1938.

51. Myers, I.b. The Myers-briggs Type Indicator : Manual Princetion, N . J. Educatronal Testing Service, 1962.

52. Peters, H.n., And Campell, G.t. Diagnosis Of Training Needs Of B-29 Mechanics From Supervisory Ratings And Self-report. Afptrc

Tech. Memorandum 55-12, 1955.

53. Roe, A. Two Rorchach Scoring Techniques : The Inception Technique And The Basic Rorchach. J. Abnor. Soc. Psychol.,. 1952, 47, 263-264.

54. Rogers, C;r. Counseling And Psychotherapy, Cambridge, Mass; Riverside, 1942.

55. Rogers, C.r. Client-centered Therapy. Cambridge, Mass. : Riverside 1951.

56. Ruggles, R. And Allport, G.w. Recent Application Of The A. S. Reaction Study. J. Abnor. Soc. Psychol., 1939, 34, 518-528.

56 A. Sundberrg, N., Assessment Of Persons N.j. Prentice Hall, Ico., 1973.

57. Sarason, S.b. The Clinical Interaction With Special Reference To The Rorchach. New York : Harper, 1954.

58. Strong, E.k. And Campbell, D. P. Manual For Strong Vocational Intrest Blank Stanford, Calif.: Stanford Univ. Press, 1966.

59. Taylor, Janet A.a. Psonnality Scale Of Manifest Anxiety. J. Normal Psychol., 1953, 48, 285-290.

60. Wallen, R.w. Clinical Psychology. New York : Mcgraw-hill, 1956.

خاتمــة

تقويم التقويم

عرضنا فى فصول هذا الكتاب المفاهيم الأساسية التى تستخدم فى الحكم على صلاحية الاختبارات وأدوات التقويم بوجه عام، ثم طبقنا هذه المفاهيم على أمثلة رئيسية من اختبارات الذكاء والقدرات والاستعدادات والتحصيل والشخصية والميول والاتجاهات.

ونعرض فى هذه الخاتمة المهام التى ترتبط بالمقوم متضمنة مهام عملية التقويم والقدرة على تقويم التقويم Meta Evaluation ونموذج يصلح لتقويم الاختبارات وادوات التقويم.

أولا : القدرة على تقويم التقويم

تتضمن القدرة على تقويم التقويم Meta Evaluation نقد وتصحيح وتصويب التعلم من الخبرات التقويمية المهمة. فالتقويم علم حديث ونحتاج الى الكثير لنتعلم منه، ومن المهم ان نشير إلى ما يلى: لا يوجد تقويم متحرر تماما من التحيز، وعلى ذلك فالمقوم الكفء هو الذى يستخدم اجراءات لمراجعة الخطط والاساليب والاجراءات وتطبيق المحكات الملائمة لرفع كفاءة وجودة التقويم.

الكفايات اللازمة للمقوم:

تتعدد الكفايات اللازمة للمقوم الجيد نعرضها فيما يلى:

١- وصف موضوع التقويم: والوصف الجيد يتضمن القدرة على التواصل الجيد مع الاخرين حول موضوع التقويم ويشمل ذلك حدود الموضوع والخصائص الهامة فيه. وتعد هذه الكفاية مهمة بصرف النظر عن موضوع التقويم (برنامج - مشروع - فكرة - اداء - مواد) فمجرد ذكر العنوان لا يكفى.

٢- وصف سياق التقويم (Context) ومعني ذلك القدرة إلى النقل الى الآخرين العوامل فى البيئة التى تؤثر فى موضوع التقويم، وفى التقويم نفسه .ويجب ان نلاحظ ان نتائج معظم التقويم التربوى خاصة ونوعية حسب الموقف، الوقت، البشر ولهذا فإن نتائج التقويم عادة ما تكون خاصة (لا تقبل التعميم) لأن العوامل فى البيئة التى تؤثر

فى الموضوع وأدائه تكون على نحو لا يقبل التكرار، ووضع هذه العوامل موضع الاعتبار يفيد فى فهم نتائج التقويم.

٣- فهم الاغراض المناسبة والاطار الملائم للتقويم: ومعنى ذلك قدرة المقوم على استخدام المعلومات الراهنة للوصول الى قرارات حول الاطار الاكثر ملائمة لتخطيط التقويم ومن ذلك مثلا:

أ- هل الموقف يحتاج ويتطلب تقويما تكوينى أم تقويما تجميعى؟

ب- هل يجب ان تكون الدراسة موجهة بالاهداف أم موجهة بالخبرة؟

ج- هل وحدة التقويم هى الفرد- الفصل - المنهج - نظام المدرسة - الإدارة ، المديرية؟

٤- تحديد واختيار الاسئلة الملائمة للتقويم والمعلومات المطلوبة ومصادر هذه المعلومات:

أى القدرة على تحديد ما نحتاج معرفته حول موضوع معين والمراد تقويمه قبل الحكم عليه، وكيف نحصل على المعلومات فتحديد ما نحتاج الى معرفته يعتمد على المحكات والمعايير والمستويات التى سوف نستخدمها فى تحديد القيمة ومن هذه المصادر.

أ- التراث العلمى المتخصص.

ب- سؤال أهل الخبرة.

ج- سؤال الطلاب - الآباء - المعلمين - فحص الملفات. أما محك أفضلية المعلومات فتعتمد على موضوعيتها وثباتها وصدقها كما يجب على المقوم تحديد الاولويات وتبريرها لجمع المعلومات حين تكون قائمة المطالب كبيرة، ويجب على المقوم معرفة أي مصدر من المصادر البديلة يهيء افضل مصدر للمعلومات الافضل مراعين فى ذلك القيود التى يعمل فيها المقوم والزمن المتاح لعملية التقويم، والتكلفة اللازمة للادوات وتجهيزها وتطبيقها وحتي الوصول الى نتائج كما يتضمن ذلك أيضا تكلفة الافراد الذين سيقومون بعملية التقويم.

٥- القدرة على تحديد واختيار الاساليب والاجراءات الملائمة لجمع المعلومات

وتجهيزها وتحليلها:

فعلى المقوم اختيار الاساليب الاكثر ملائمة من بين الاساليب المتعددة المتاحة مثل الاختبارات – مقاييس التقدير – الاستبيانات – المقابلات – قوائم الملاحظة كما عليه ان يختار المنهج الملائم من التصميمات التجريبية أو طرق المسح مع القدرة على تسجيل وتجهيز الانواع المختلفة من المعلومات مثل تلخيص البيانات الكمية وتحليلها احصائيا والقدرة ايضا على تنظيم البيانات الكيفية وهذه المهارات الفنية تتطلب تحديدا واختبار افضل منحى مع القدرة على اجراءاها.

٦– تحديد قيمة موضوع التقويم أو الحكم عليه:

ويعنى ذلك القدرة على تطبيق محكات على المعلومات الوصفية حول الموضوع للوصول الى عبارة قيمية يمكن الدفاع عنها وتوجد طرق عديذة لتحديد القيمة ومهمة المقوم ان يكون قادرا على التقويم على نحو منتظم وصولا الى نقطة الاعداد للأحكام أو عبارات القيمة.

٧– نقل خطط التقويم ونتائجه:

إن القدرة على فهم الحاجات المعلوماتية لدى الجمهور العام وإعداد وسائل مناسبة وتوزيع كل رسالة على نحو يمكن لها ان يستمع اليها، فالوثائق المكتوبة والتقارير الفنية الطويلة عادة ما تفوقها تلك التى يقدمها المقوم ولكنها ليست بالضرورة افضل الطرق لنقل الخطط والنتائج. فهذه التقارير الفنية قد تفيد كوثائق تتاح لاي شخص يرغب فى متابعة نقطة معينة وعلى ذلك فالمقوم مطلوب منه معرفة وفهم جمهور المستمعين لمعلومات التقويم وينتقى افضل طريقة لنقل المعلومات اليه بطريقة تجذب الانتباه والقيام بطريقة مفهومه ولا تستغرق وقتا طويلا.. ومتابعة التواصل مع الجمهور غالبا ما يكون جزءا من عملية التواصل.

٨– القدرة على ادارة التقويم:

أ– أى القدرة على التخطيط لانشطة التقويم.

ب– تحديد الموارد البشرية والمالية لإجراء مهام التقويم.

ج– توافر قيادة طوال الدراسة.

د- الدعم والمراقبة والاشراف على الافراد للوصول إلى تقويم على مستوى رفيع والمقوم الجيد هو الذى يستطيع أن يستثمر الموارد المتاحة لانتاج جدول للتقويم الجيد بحيث يشمل جميع العناصر ويتطلب ذلك:

أ- القدرة على توجيه المهام

ب- الخبرة والمهارة فى العمل مع الناس.

جـ- مهارات فى التخطيط

د- مهارات اتخاذ القرار.

٩- الالتزام والتمسك بالميثاق الخلقى:

أ- اى القدرة على الالتزام بالسلوك المهنى طوال جوانب التقويم

ب- معرفة قواعد واصول حماية الانسان وسرية المعلومات.

ج- عند التعامل مع معلومات أو إصدار الاحكام القيمية على نحو يؤثر فى العمل ورفاهية الاخرين يجب على المقوم الالتزام بالمعايير الاخلاقية فمعرفة الالتزام بالمباديء الاخلاقية هو جزء لا يتجزأ من السلوك المهنى.

١٠- التكيف للعوامل الخارجية التى تؤثر فى التقويم ويتطلب ذلك:

أ- المرونة خلال عملية التقويم.

ب- تمثل المعلومات الجديدة فى اجراءات التقويم بحيث تصبح هذه المعلومات معلومة.

ج- القيود التى تؤثر حتي على الخطط الأكثر تنظيما وتشمل ما هو قانونى - تنظيمى - سياسى- ادارى - انسانى - منهجى وعلى المقوم ان يكون دجماتيقيا فى تنفيذ خطة التقويم والخلاصة ان التقويم يتضمن جوانب واساسيات مهنية واخلاقية واجتماعية

ثانيا: تقويم ادوات التقويم:

نظرا لما يتطلبه التقويم من دقة وجودة فيما يتضمنه من اعداد وتطبيق واصدار حكم فإن من الضرورى التأكيد على جودة أدوات التقويم. وفيما يلى نعرض لنموذج لتقويم أدوات التقويم.

يتألف النموذج المقترح من ٢٨ بندا، وفيه تتطلب البنود رقم ١ إلى رقم ٩ وصفا بسيطا للاختبار يمكن أن تحصل عليه عادة من كراسة تعليماته والدليل الخاص به. ويتطلب البند رقم ٦ وصف الطابع العام للاختبار، كما أن البند رقم ٧ (تاريخ النشر) ليس على درجة كبيرة من الأهمية لأن بعض الاختبارات القديمة قد تكون ممتازة، وأى اختبار قديم يجب أن يفحص بعناية ودقة، فقد تكون بعض أسئلته غير ملائمة لتغير العصر، وقد تكون معاييره غير مفيدة، وقد يكون دليله غير كامل، وفى هذا البند يجب أن نسجل أى طبعات جديدة أو أى تعديلات طرأت على الاختبارات.

ويحتوى النموذج المشار إليه على البند رقم ١١ الذى يتطلب وصف الاختبار من حيث عناصره وأسئلته وتصحيحه من أجل تكوين انطباع عام عن الاختبار ولاشك أن هذا البند هو الذى يحدد مدى جاذبية الاختبار للمفحوص وللرجل العادى، أما البند ١٢ يتطلب تحديد مقاصد المؤلف سواء كانت انتقاء أو توجيها أو علاجا أو تقويما مدرسيا، وبالتالى يحدد بهدف المؤلف إلى قياسه بالاختبار، وقد تتضمن بعض كراسات التعليمات الأسلوب الإحصائى الذى استخدمه المؤلف فى انتقاء الأسئلة أوتحليلها وحساب صدق تلك الأسئلة.

ويتضمن البند ١٣ حكما على تعليمات الاختبار فى ضوء الأسس العامة، التى عرضناها فى الفصل الثانى من الكتاب، أما البند ١٤ فيتطلب أن نقوم بتحليل نظرى لأسئلة الاختبار للحكم على ما تتضمنه من قدرات أو خبرات، أو عادات أو سمات شخصية، وكيف تؤثر هذه العوامل كلها أو بعضها فى درجة الاختبار، وهذا التحليل ضرورى لكل درجة من درجات الاختبارات الفرعية التى يحتويها الاختبار وبالطبع نحن لانستطيع أن نحدد جميع العوامل المسهمة فى درجة الاختبار، ولكن مثل هذا العمل التحليلى يثير عددا من الأسئلة يمكن أن تستخدم فى تفسير نتائج الاختبار، ويجب أن يسجل التقرير ما تدل عليه الدرجة فى الاختبار سواء كان حكمنا هذا يتفق مع مقصد المؤلف أو لا يتفق ، كما يوضح العوامل غير المرتبطة التى تؤثر فى الدرجة وقد تفسدها، ويمكن للباحث أن يلجأ إلى تطبيق الاختبار على نفسه أو على مفحوص آخر لديه خبرة بالقياس النفسى ثم يلاحظ الباحث أداءه فى الاختبار، أما فى حالة اختبار المفحوص لنفسه فإن الملاحظة تصبح من نوع التأمل الباطنى intro-spection، ورغم الاعتراضات المنهجية على هذه الطريقة إلا أنها تفيد كثيرا فى

إعطاء كثير من الدلالات عن طبيعة الاختبار.

وتدل البنود من ١٦ – ١٩ ما يتوفر عن صدق الاختبار من بيانات تجريبية، وذلك للحكم على مدى صلاحية هذه البيانات لأغراض الاختبار، أما البنود من ٢٠ – ٢٢ فتشمل بيانات ثبات الاختبار، ويتصل للبند ٣٢ بثبات الاختبار من حيث فاحصه ومصححيه، والمعايير من حيث دقة تفسير الأداء فى الاختبار كما يدل عليها البندان ٢٤، ٢٥.

ويشير البند ٢٦ إلى تلخيص العروض الناقدة للاختبار فى المؤلفات والمجلات المتخصصة، أما البند ٢٧ وموضوعه «تقويم عام للاختبار» فيتضمن تلخيصاً عاما نهائيا لنواحى الضعف والقوة مع اعتبارالنواحى العلمية والفنية الوظائف العامة، بالإضافة إلى تسجيل الملاحظات العامة للفاحصين والمفحوصين وتعليقاتهم الخاصة عليه.

وهكذا يتوافر للفاحص قائمة تحليلية للاختبار بحيث يستطيع الفاحصون - وخاصة أولئك المشاركون فى البرامج الشاملة للتقويم سواء التربية أو الصناعة أو القوات المسلحة - أن يختاروا الاختبار المناسب للغرض المناسب فى الوقت المناسب، وبأقل فاقد ممكن فى الجهد والوقت والمال.

نموذج لتقويم الاختبارات

كما ذكرنا أن هذا النموذج يتكون من البنود التالية:

١ – عنوان الاختبار.

٢ – اسم مؤلف الاختبار ومترجمه (إن وجد).

٣ – الناشر.

٤ – صور الاختبار والجماعات التى يمكن أن تطبق عليها.

٥ – المعالم العلمية (أية خصائص عملية تتصل بالأداء فيه أو تصحيحه).

٦ – نوع الاختبار (اختبار قدرات أو شخصية مثلا).

٧ – تاريخ النشر.

٨ – التكاليف ، كراسة التعليمات والأسئلة، ورق الإجابة.

٩ – الزمن المطلوب.

١٠ – الغرض الذى من أجله نحكم على الاختبار.

١١ – وصف الاختبار من حيث عناصره، وتصحيحه.

١٢ – غرض المؤلف وأساس اختبار عناصر الاختبار عنده.

١٣ – دقة التعليمات ووضوحها، والتدريب المطلوب لمن يقوم بعملية إعطاء الاختبار، ودرجة تقنينها.

١٤ – الوظائف أو السمات العقلية الممثلة فى كل درجة.

١٥ – تعليقات وملاحظات عن تصميم الاختبار وبنائه (ورقة الأسئلة وطريقة الإجابة).

١٦ – الصدق التنبؤى (المحك، وعدد ونوع الحالات التى استخدمت فى حسابه، والنتيجة).

١٧ – الصدق التلازمى (المحك، وعدد ونوع الحالات المستخدمة ،والنتيجة).

١٨ – أية أدلة تجريبية أخرى تبين ما يقيسه الاختبار (صدق التكوين الفرضى).

١٩ – تعليقات وملاحظات عن صدق الاختبار بالنسبة لأغراض معينة سيستعمل من أجلها الاختبار.

٢٠ – ثبات الاختبار بحساب معامل الاستقرار (الطريقة التى اتبعت للوصول إليه

وعدد ونوع الحالات ، والنتيجة .

٢١ – ثبات الاختبار بحساب معامل التكافؤ (الطريقة التى اتبعت للوصول إليه، وعدد نوع الحالات، والنتيجة) .

٢٢ – ثبات الاختبار بحساب معاملات الاتساق الداخلى أو اتساق الأسئلة (الطريقة التى اتبعت للوصول إليها، وعدد ونوع الحالات والنتيجة) .

٢٣ – موضوعية الاختبار من حيث التصحيح وتفسير الدرجات.

٢٤ – المعايير (نوعها، وعدد ونوع الحالات) .

٢٥ – تعليقات وملاحظات عن دقة فئات الاختبار وصحة المعايير للغرض الذى سيستعمل الاختبار من أجله .

٢٦ – تقويم وملاحظات من قاموا باستعراض للاختبار وتقويمه ونقده .

٢٧ – تقويم عام للاختبار.

٢٨ – المراجع.

هذا الكتاب

الإنسان هو محور التنمية والتقدم ، والاهتمام به أصبح من متطلبات العصر بما فيه من تسارع ومنافسة فى إثبات الذات واستثمار مقوماتها . ولايكون ذلك إلا من معرفة الذات سواء فى قوة مقوماتها فتستثمر وتنمو، وتلك التى تحتاج إلى رعاية وتوجيه .

وبالطبع لا يتم ذلك إلا من خلال التقويم الذى هو بؤرة الاهتمام فى هذا الكتاب سواء لمعرفة أسس عملية التقويم والقياس ، وأدواتهما ، وشروط استخدامها وتفسير نتائجها ، وتوجيه ذلك فى استثمار قدرات الإنسان .

المؤلفون

ISBN 977-05-0529-3

9 789770 505298

مكتبة الأنجلو المصرية
THE ANGLO-EGYPTIAN BOOKSHOP
The World of Words & Thoughts
www.anglo-egyptian.com